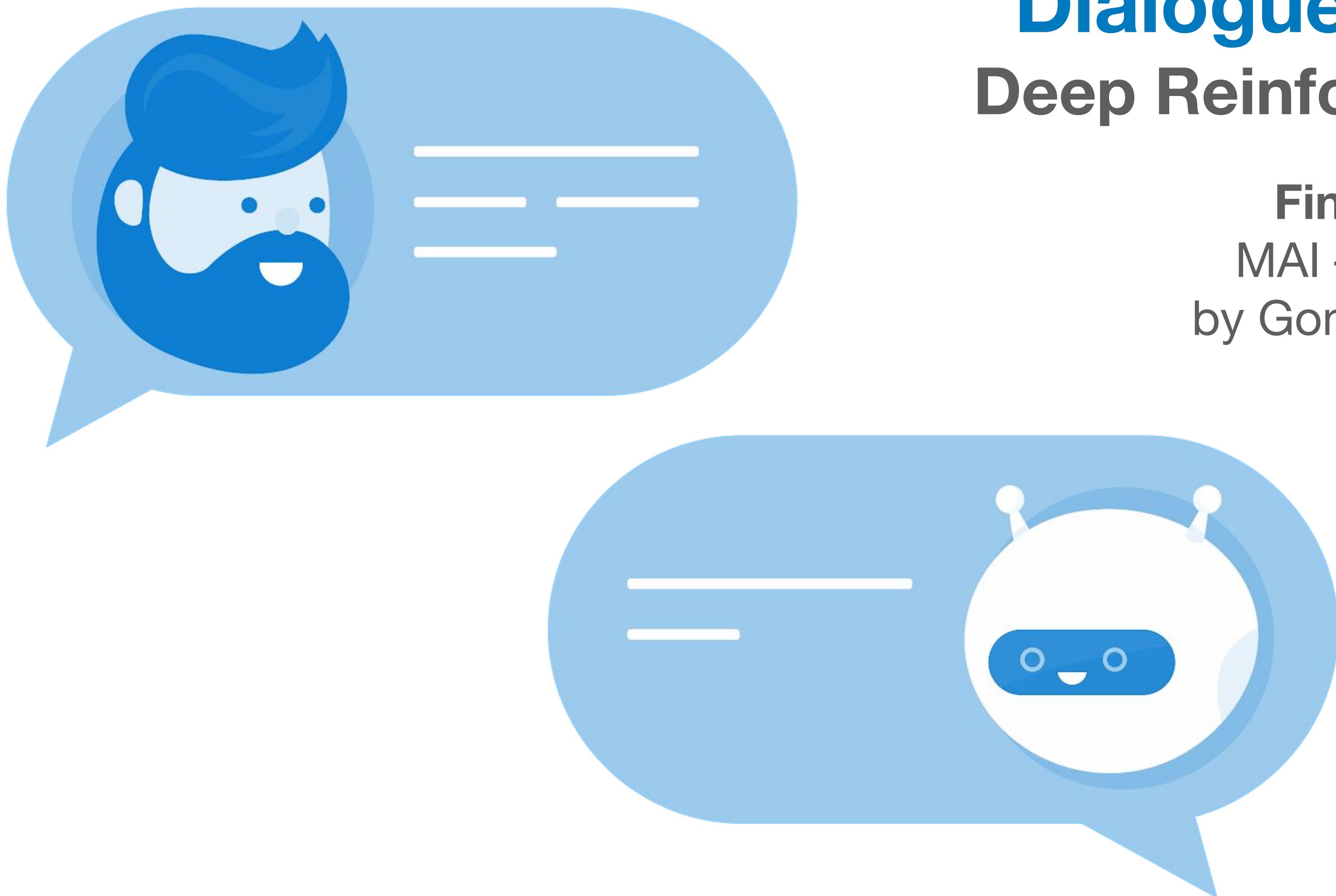


# Dialogue Generation Chatbot

## Deep Reinforcement Learning chatbot



Final Project Presentation

MAI – CIR January 18th 20201

by Gonzalo Recio & Jana Reventós

# Presentation content

- 1. Motivation**
- 2. Project proposal**
- 3. State of the art**
- 4. Solution proposal**
- 5. Model experiments**
- 6. Human-interaction experiments**
- 7. Conclusion & Future work**

# Motivation

- A chatbot is a system, powered by rules and AI, which **simulates a real interaction with users** via chat interface.
- Closed-domain chatbots respond to keywords or intents to accomplish a specific task, **open-domain chatbots** can engage in conversation in any topic.
- **State-of-the-art language models** (e.g. BERT & GPT) have shown **outstanding results** in NLP tasks.
- Tend to easily **generate incoherent sentences** or highly generic responses such as “I don’t know” regardless of the input on the long term (Li J. et al. 2016), therefore, **lack of engagement**.



# Project proposal

Implement a conversational voice-based chatbot using **reinforcement learning**. Improve users engagement by making **dialogues last long** to have a fluid and natural conversation.

Create a **friendly interface** to make users feel **comfortable**.

## MAIN OBJECTIVE

Develop and design a voice-based chatbot with the aim of strengthening user engagement during the interaction experience.

# Project proposal | Objectives



## Objective 01.

**Develop** a dialog generation chatbot using RL

- RL train by dialoguing to improve long-term reward

## Objective 02.

Build up a prototype of physical chatbot assistant.

## Objective 03.

Test the interactive behavior of the chatbot with real users.

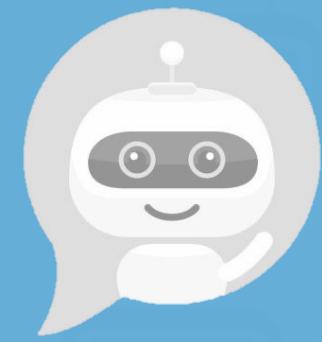
## Objective 04.

Evaluate the model and interaction with users.

# Project proposal | Objectives

## Objective 02.

Develop a dialog generation chatbot using RL.



## Objective 02.

**Build** up a prototype of physical chatbot assistant.

- Create a human – friendly interface with sentiment analysis pre-trained model

## Objective 03.

Test the interactive behavior of the chatbot with real users.

## Objective 04.

Evaluate the model and interaction with users.

# Project proposal | Objectives

## Objective 02.

Develop a dialog generation chatbot using RL.

## Objective 02.

Build up a prototype of physical chatbot assistant.



## Objective 03.

**Test the interactive behavior of the chatbot with real users.**

- Interaction experiment design
- Experiment with real users

## Objective 04.

Evaluate the model and interaction with users.

# Project proposal | Objectives

## Objective 02.

Develop a dialog generation chatbot using RL.

## Objective 02.

Build up a prototype of physical chatbot assistant.

## Objective 03.

Test the interactive behavior of the chatbot with real users.



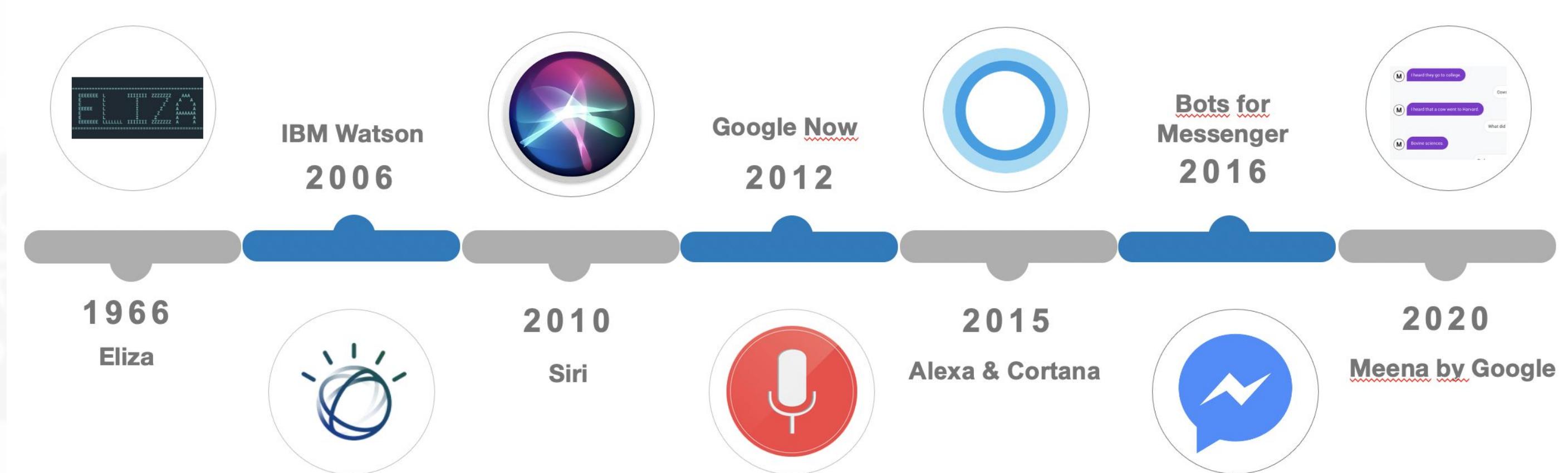
## Objective 04.

### **Evaluate model and interaction with users.**

- Model and Human evaluation
- Hypothesis validation

# State – of – the – art | Conversational dialogue systems

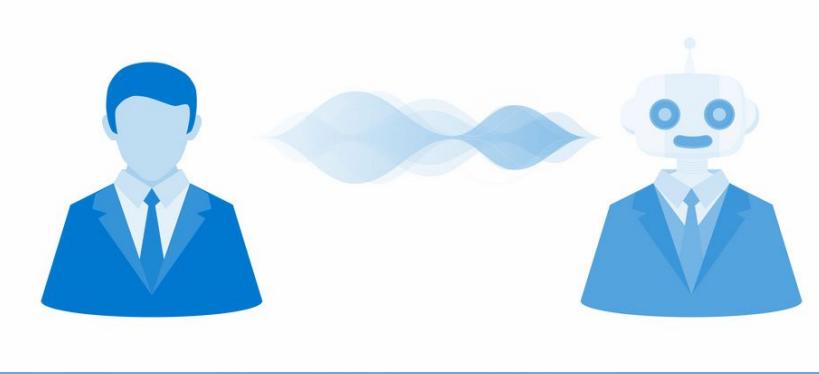
- **End-to end neural network approaches** (Shang et al., 2015; Vinyals and Le, 2015; Sordoni et al., 2015; Serban et al., 2016; Zhang et al., 2019; Adiwardana et al. 2020) offer the **simplicity of a single learned model**. The state of the art use **GPT-2 transformer** model and **evolved transformer** (So et al. 2019) **with seq2seq model**.
- Integration of seq2seq models (encoder-decoder RNN) with **reinforcement learning (RL)** paradigms. RL improves dialogue long-term performance (Li J. et al. 2016).



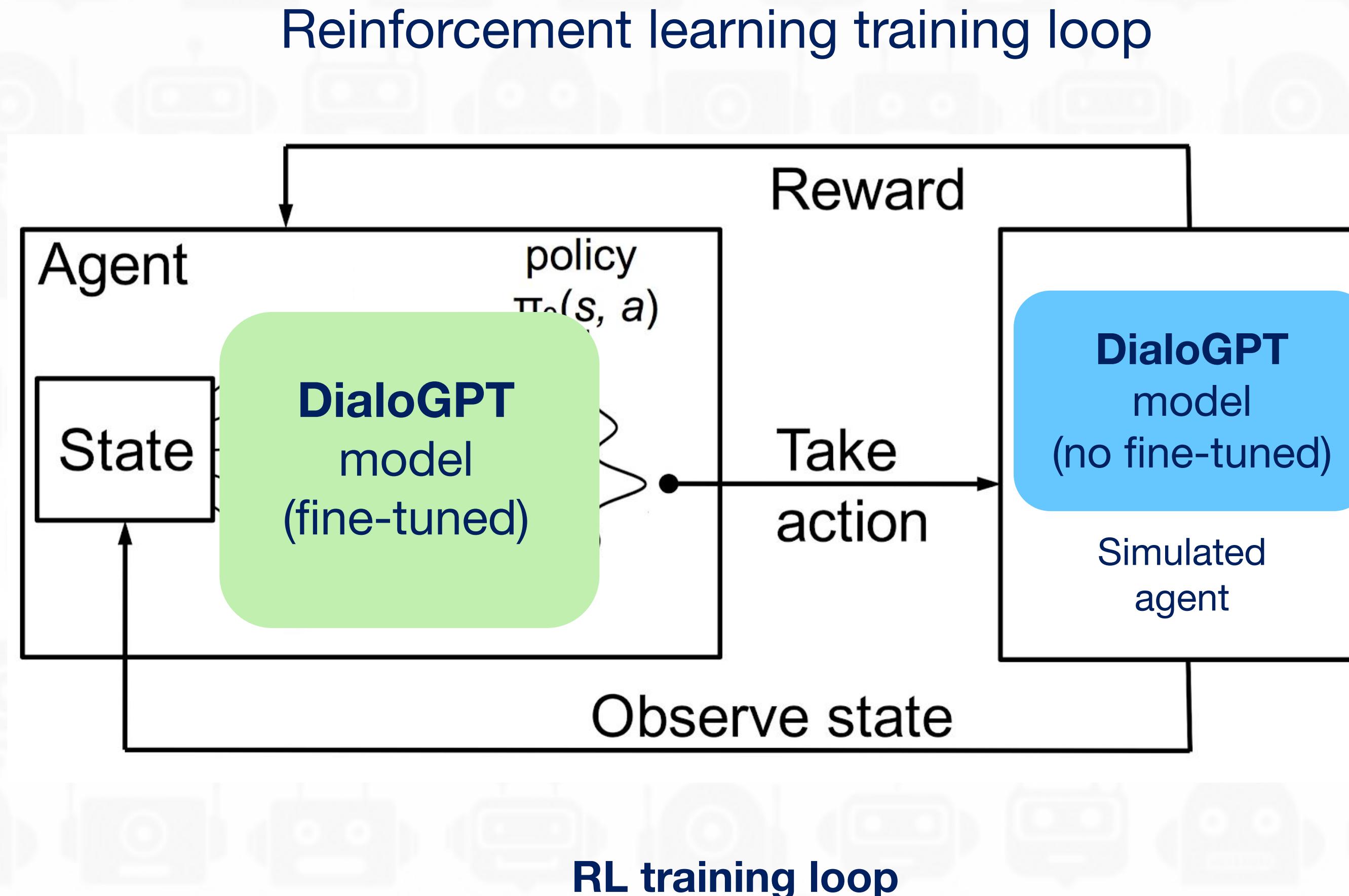
1. <https://www.msxiaobing.com/>
2. <https://www.pandorabots.com/mitsuku/>
3. <https://www.cleverbot.com/>

# Solution proposal | Cognitive frame

## Reinforcement learning for open-domain dialogue

AGENTS	ACTION	STATE	POLICY	REWARDS
 - <b>Chatbot agent:</b> sentences generated by the fine-tuned DialoGPT model ( $c_i$ ) - <b>Simulated agent:</b> sentences generated by base the DialoGPT model ( $u_i$ )	 The action is the <b>dialogue utterance</b> generated by the fine-tuned DialoGPT model. The action space is infinite.	 The previous dialogue turns $[(c_1, u_1), (c_2, u_2), \dots, (c_{i-1}, u_{i-1})]$ Called <b>chat history</b> .	$f(x)$ Defined by the <b>DialoGPT language model</b> . Mutual Information maximization.	 1. Ease of answering 2. Information flow 3. Repetition avoidance

# Solution proposal | Cognitive frame



# Solution proposal | Cognitive frame

## Rewards

- **Ease of answering ( $r_1$ ):** a turn generated by the chatbot should be ease to respond to, avoiding dull responses.
- **Information flow ( $r_2$ ):** chatbot able to contribute new information to keep the dialogue moving and avoid repetitive response.
- **Repetition avoidance ( $r_3$ ):** we want the chatbot to avoid repeating the sentences of the other user.

$$r_1 = -\frac{1}{N_S} \sum_{s \in S} \frac{1}{N_s} \log p_{\text{seq2seq}}(s|a) \quad r_2 = -\log \cos(h_{p_i}, h_{p_{i+1}}) = -\log \cos \frac{h_{p_i} \cdot h_{p_{i+1}}}{\|h_{p_i}\| \|h_{p_{i+1}}\|} \quad r_3 = -\log \cos(u_i, c_i)$$

$$r(a, [u_i, c_i]) = \lambda_1 r_1 + \lambda_2 r_2 + \lambda_3 r_3$$

Instantaneous reward

$$R_t = \sum_{i=t}^T \gamma^{(i-t)} r(a, [u_i, c_i])$$

Expected reward

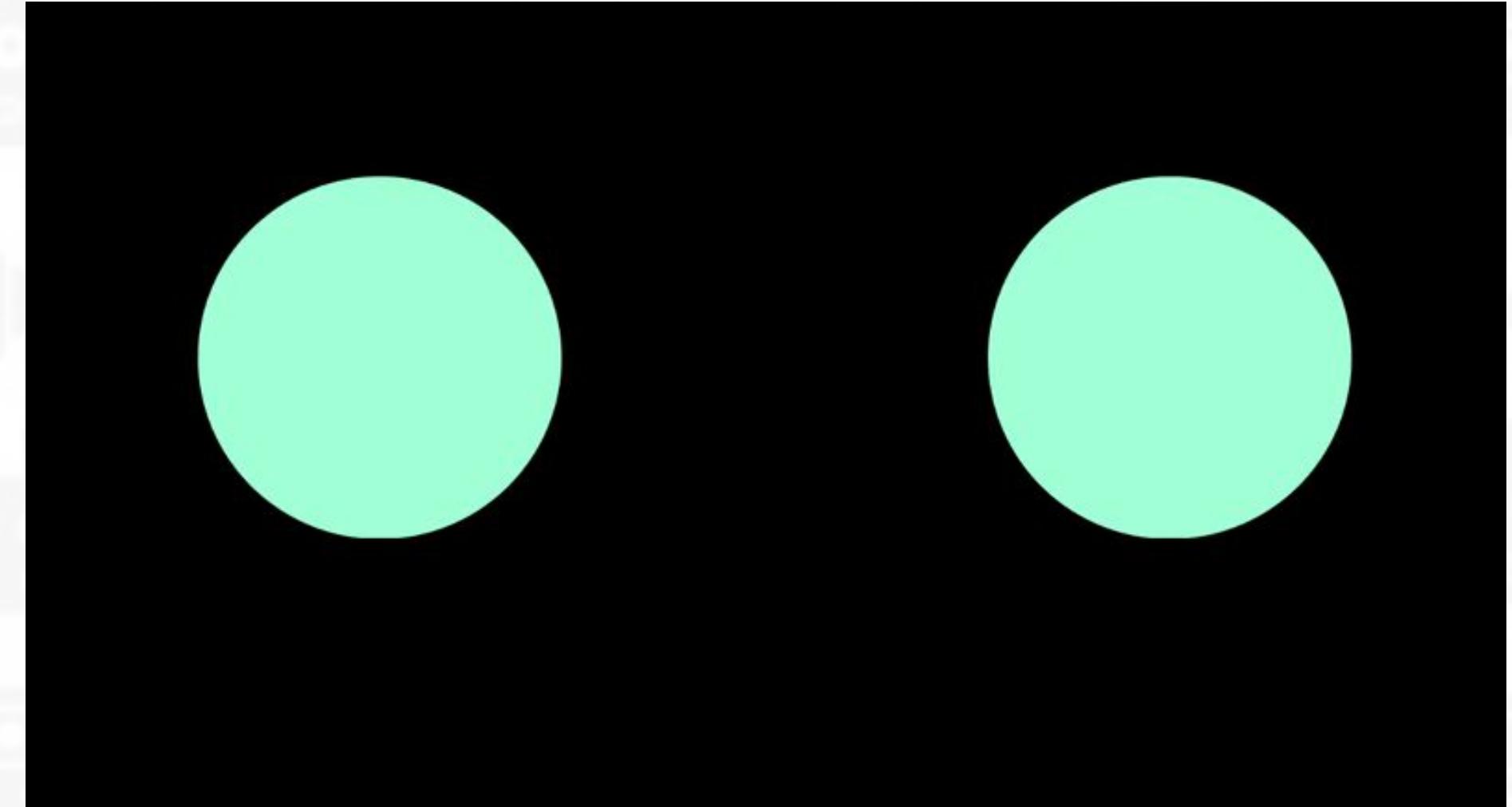
# Solution proposal | User interaction

## PROBLEM

Chatbots usually have inadequate conversational user interfaces, especially for old people or people without experience with this technology.

- User centered-design
- Build a simple conversational interface
- Create a bot personality

We a user interface based on chatbot **eyes expression** with **sentiment analysis**.



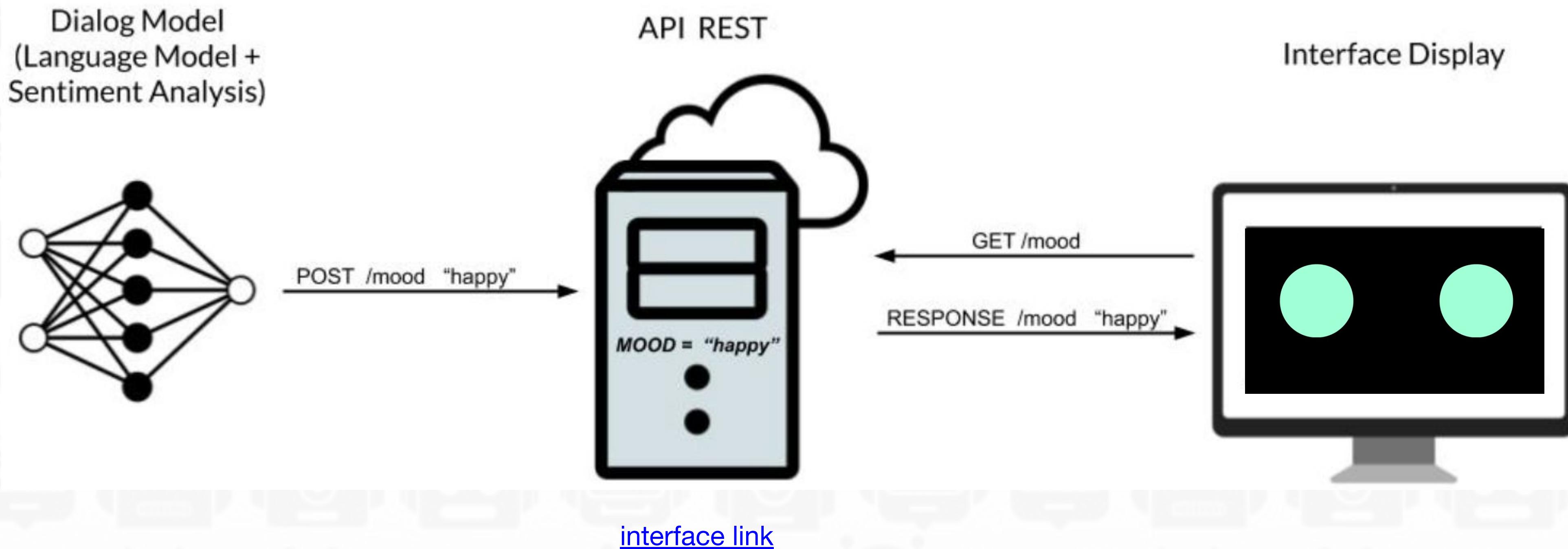
[LINK TO CHATBOT USER INTERFACE!](#)



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH

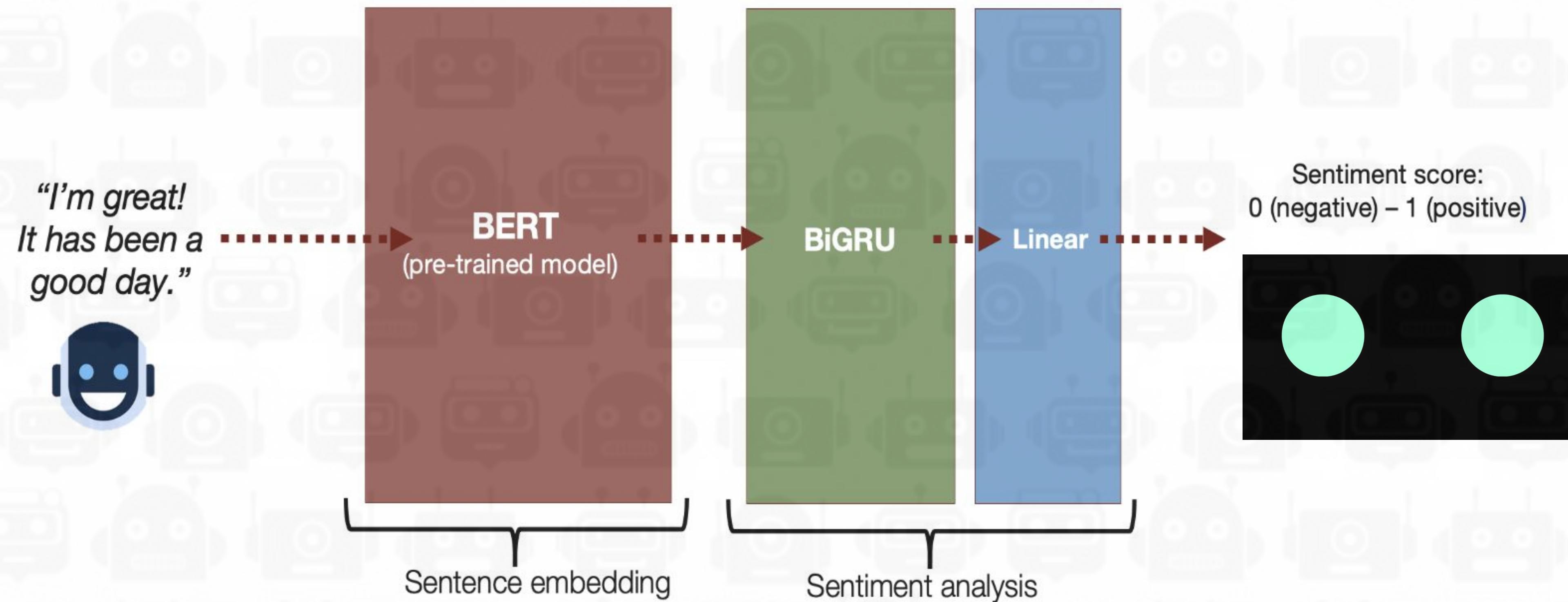
# Solution proposal | User interaction

## Chatbot GUI: Decoupling the model from the interface

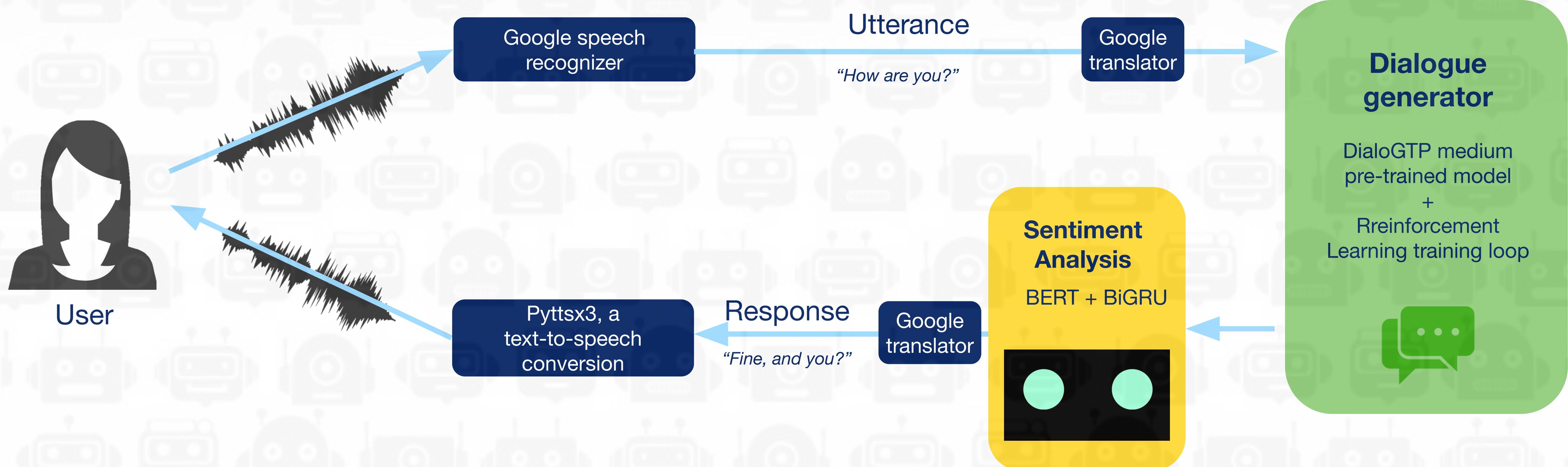


# Solution proposal | User interaction

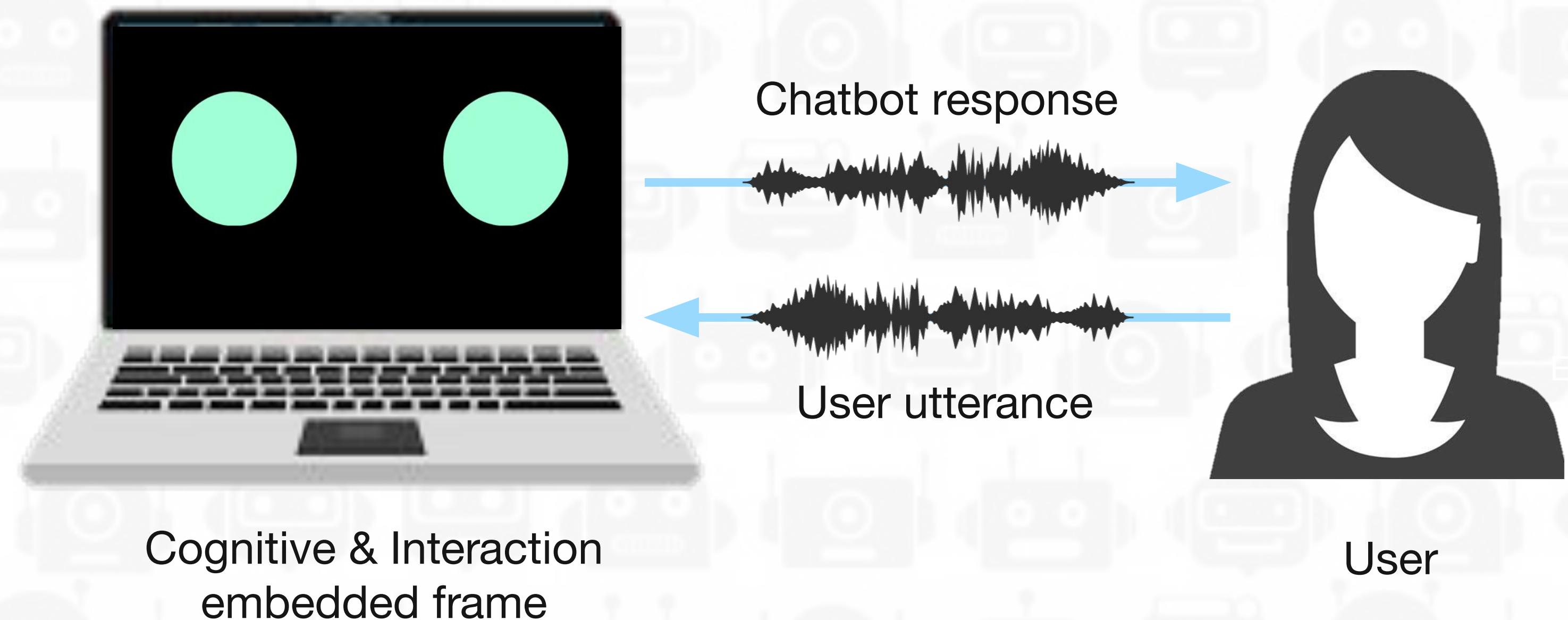
## Sentiment Analysis



# Solution proposal | Model architecture



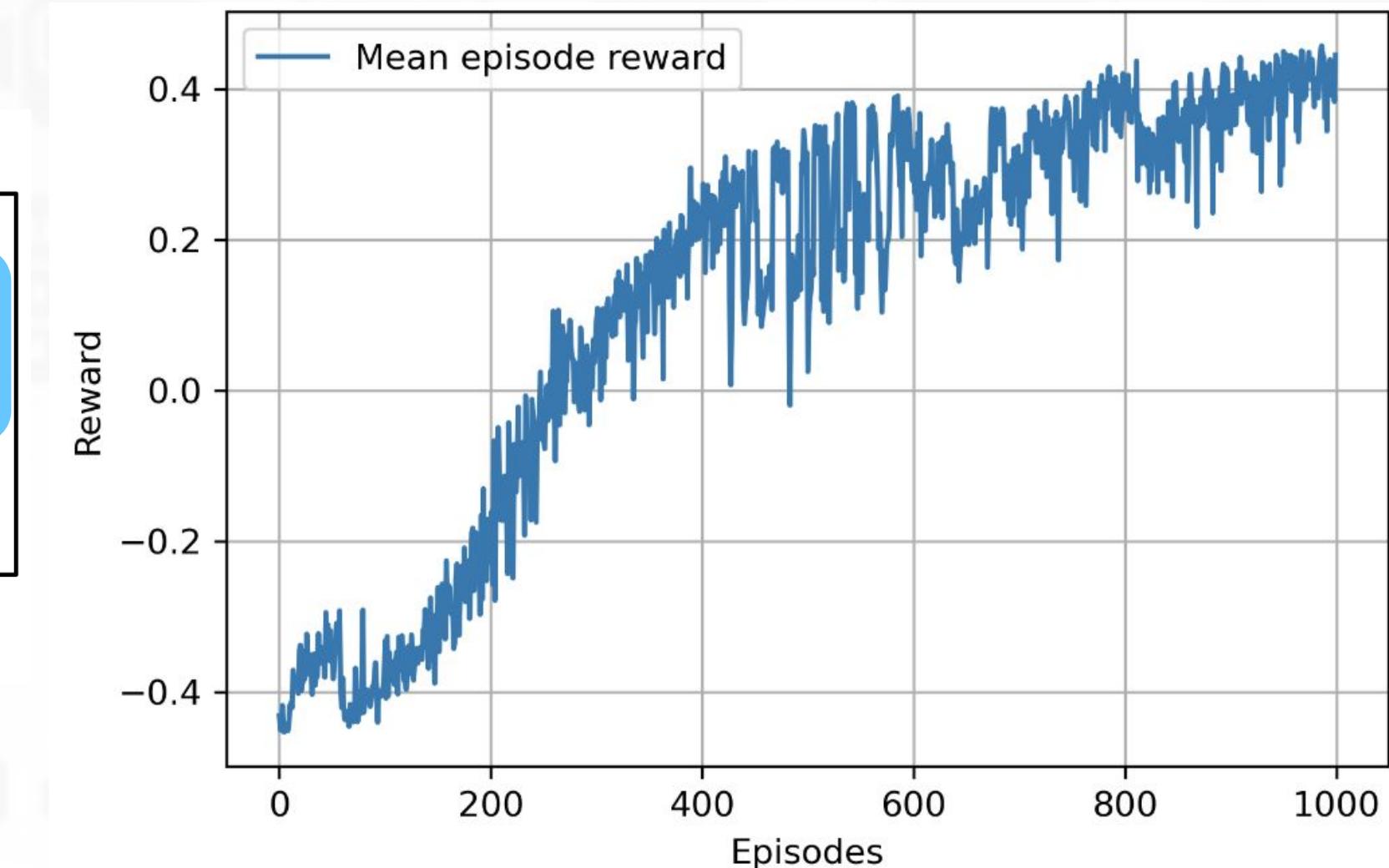
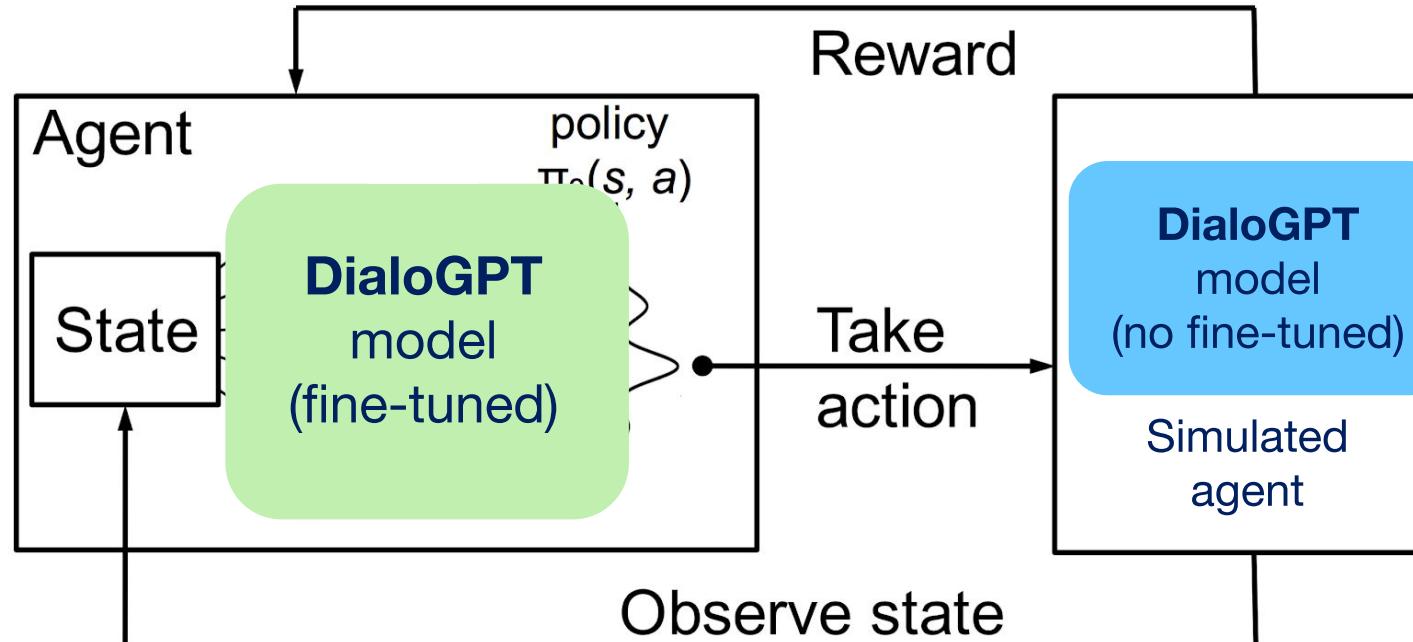
# Solution proposal | Prototype setup



# Model experiments | DialoGPT model with RL training

## Quantitative evaluation

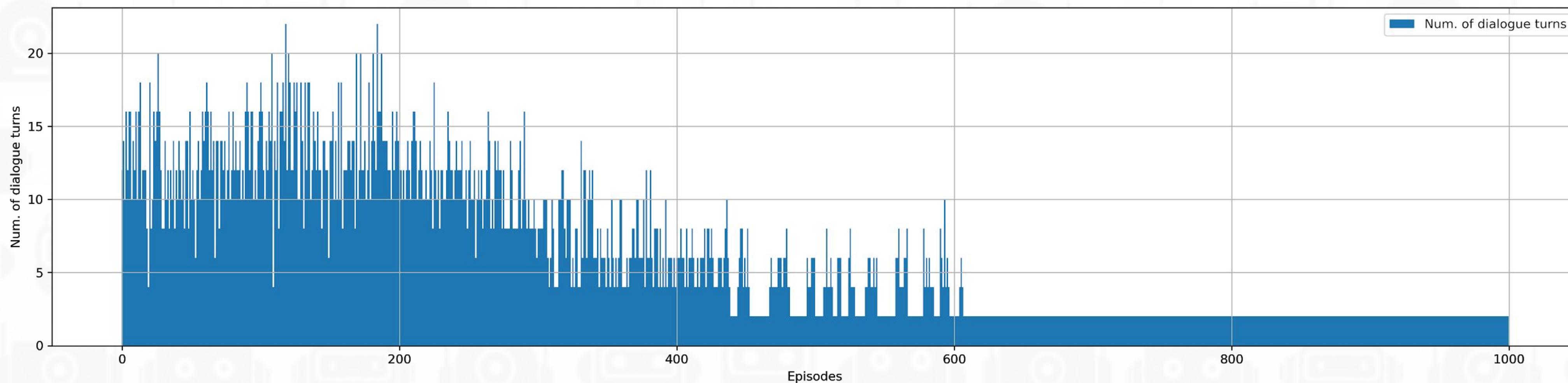
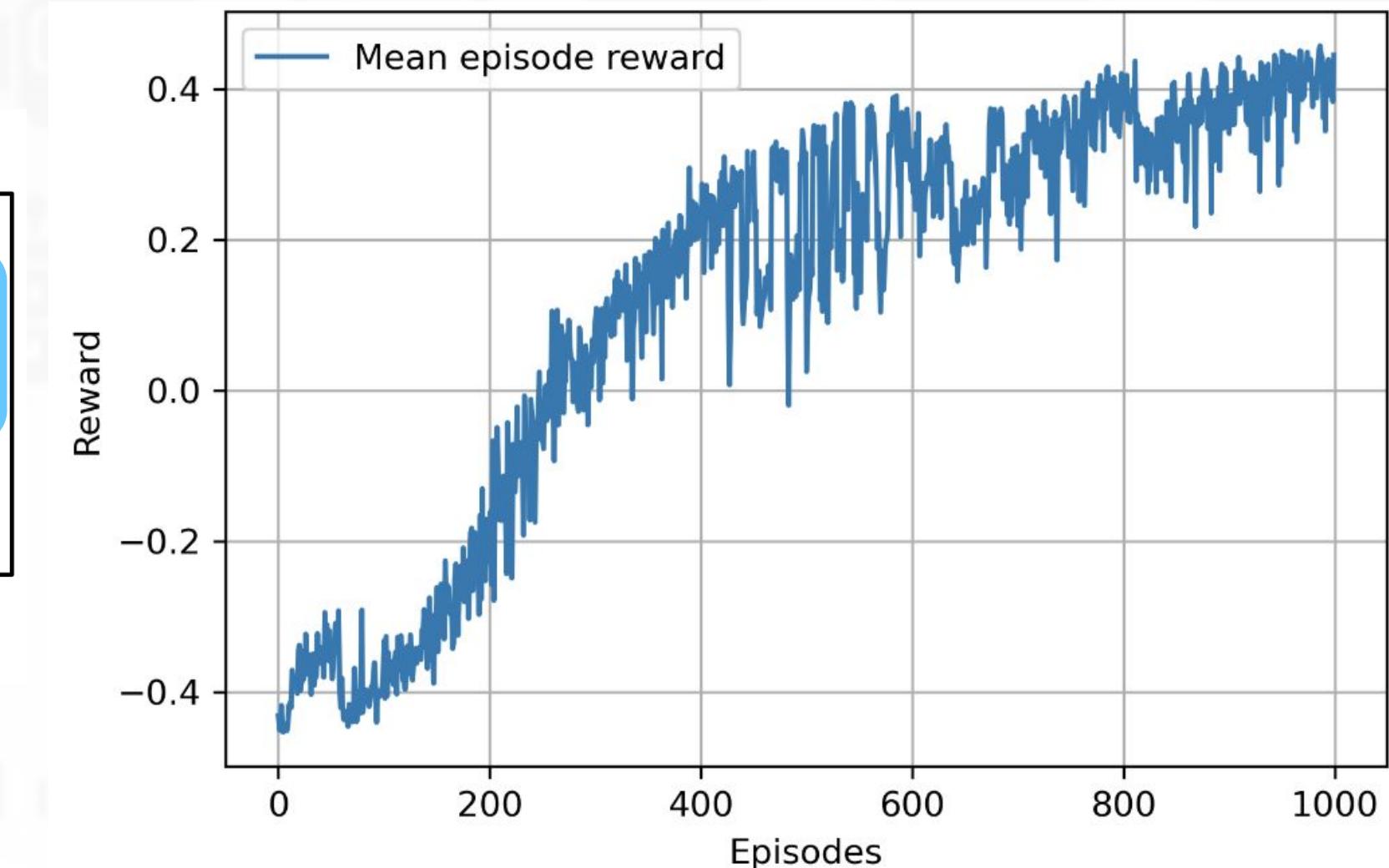
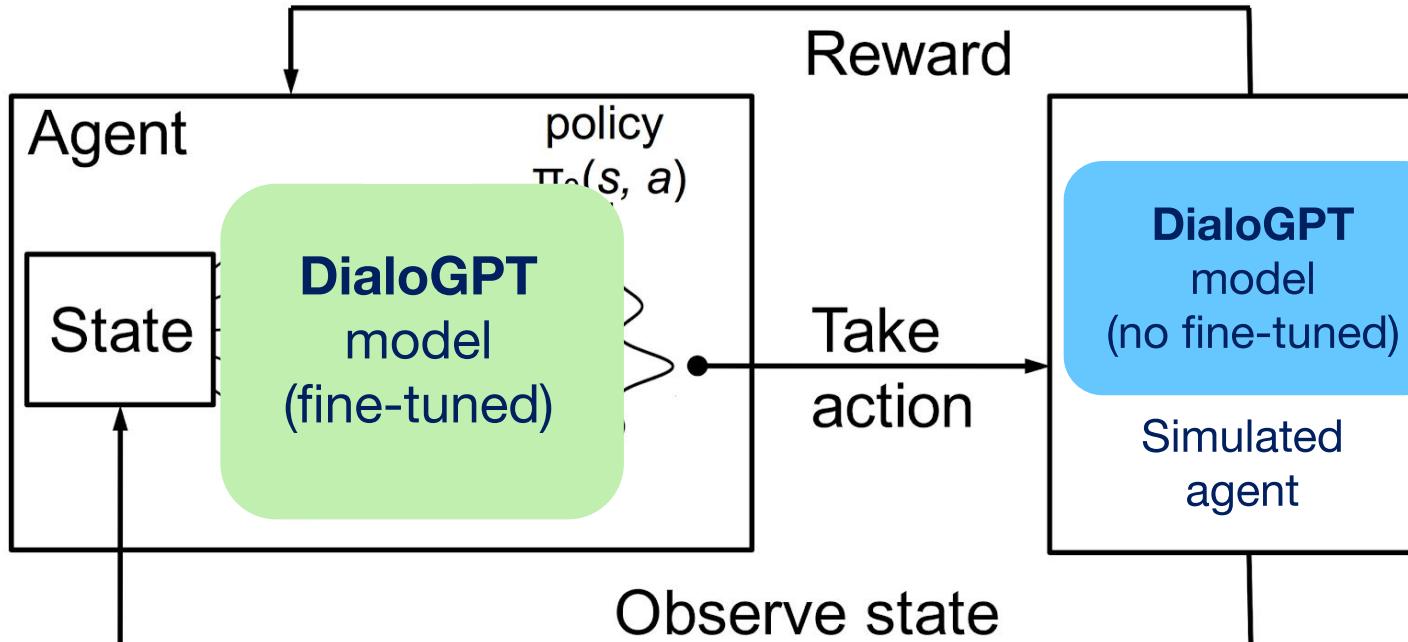
- Track of the achieved reward  
$$r = 0.5 r_1 + 0.25 r_2 + 0.25 r_3$$
- Generated dialogue turns



# Model experiments | DialoGPT model with RL training

## Quantitative evaluation

- Track of the achieved reward  
 $r = 0.5 r_1 + 0.25 r_2 + 0.25 r_3$
- Generated dialogue turns



# Model experiments | DialoGPT model with RL training

# Qualitatively evaluation

- o Evaluate the model's ease of answering observing the evolution of simulated dialogues along episodes.

## Episode 1:

User: Hello  
**DialoGPT**: Hello! :D  
User: How are you?  
**DialoGPT**: I'm good!  
User: That's good!  
**DialoGPT**: I'm good too!  
User: That's great!  
**DialoGPT**: So, what's up?  
User: Nothing much.  
**DialoGPT**: That's good.  
User: That's great.  
**DialoGPT**: I'm good  
User: That's good  
**DialoGPT**: I'm good.  
User: That is good  
**DialoGPT**: I'm  
User: I'm good, too.  
**DialoGPT**: I'm  
User: Good  
**DialoGPT**: I

## Episode 200:

User: Hello  
**DialoGPT:** Hi there!  
User: How are you?  
**DialoGPT:** Fine, how are you?  
User: I'm good, how are you  
**DialoGPT:** I'm good too, thanks. How are you?  
User: Good, how are you doing?  
**DialoGPT:** I'm doing well, well, well, how are you going?  
User: I'm going to be busy for a while, but I'm going to be doing a lot of work.  
**DialoGPT:** Ah, well, that's fine then. I'm sure you'll be busy  
User:  
**DialoGPT:**

Episode 400

User: Hello

**DialoGPT:** Hey there! How are you doing today?

User: I'm doing well, how about you?

**DialoGPT:** doing well too! How bout you?

User: I'm good, how about you

**DialoGPT:** doing well too thanks for asking me questions about stuff about myself that i've never heard of before today huhuhuhuhuhuhuhuhuhuhuhuhueueueueueueueuhuhueueuehueueueuhueueueuhuhuhuhueueueueueuuueue

User:

**DialoGPT:**

Episode 800

User: Hello

**DialoGPT:** Hey buddieinoeus-  
aurusesaurusesaurusesauru  
saurusesaurusesauriaurus-  
esaurusesauraaurusesauruses  
aurosaurusesaurusesaunaau-  
rusesaurusesaunosaurusauru  
sesaurusesauseaurusesaurus-  
esaulosaurusaurusesaurule  
saurusesaurusesausaurusaurus-  
esaurulesaurusaurusesau  
ruitsaurusesaurusesaultause-  
aurusesauruitsauseauruse  
saurulesauseaurusesaurue-  
saurusesaurusesaveauseaurus-  
esaur

User:

**DialoGPT:**

# Model experiments | DialoGPT model with RL training

## Models comparison

- **DialoGPT-medium** (0 episodes) vs **DialoGPT-medium RL** (with 60 episodes)
  - Average turns
  - Average duration
  - Average unigrams and bigrams

Results with real users:

Model	Avg. turns	Avg. duration (s)	Avg. unigrams	Avg. bigrams
DialoGPT standard	15.8	257.693	69.7	253.3
DialoGPT with RL	<b>31.7</b>	<b>605.251</b>	<b>253.3</b>	<b>337.5</b>

Comparison between the baseline model and the RL trained model. Results have been averaged among dialogues performed by the same model.

# Model experiments | DialoGPT model with RL training

User: hola  
Chatbot: ¡Hola! :D  
Chatbot: Could not understand audio  
User: None  
Chatbot: Eres el mejor  
rUser: qué  
Chatbot: ¿Qué?  
User: qué porque soy el mejor  
Chatbot: Ah, vale.  
User: quién eres tu  
Chatbot: No sé quién eres, pero probablemente eres mejor que yo.  
User: soy un humano  
Chatbot: Eso es una mentira.  
User: como te atreves  
Chatbot: ¡No eres humano!  
User: entonces que soy  
Chatbot: ¿Un humano?  
User: lo estás dudando  
Chatbot: No soy.  
User: te gusta el fútbol  
Chatbot: Hago.  
User: que si prácticas el fútbol  
Chatbot: Eso es verdad.  
User: entonces conoces a messi  
Chatbot: Sí, es un buen jugador.  
User: estamos de acuerdo  
Chatbot: Todos estamos de acuerdo  
User: es que es el mejor  
Chatbot: El es el mas grande  
User: y te gusta el cine  
Chatbot: ¡Amo la película!  
User: qu' e pel'cula te gusta  
Chatbot: El de los mejores actores y actrices  
User: None  
Chatbot: no se que decir  
User: una pregunta muy difícil  
Chatbot: Es una vida dura  
User: la de los actores y la de los robots  
Chatbot: Esa es una pregunta difícil.  
User: pero t' u eres muy lista  
Chatbot: No estoy seguro de si debería estar orgulloso o avergonzado por este comentario.  
User: muy orgullosa  
Chatbot: Muy orgulloso  
User: None  
Chatbot: Estás bien, solo estoy tratando de hacerlo más fácil para mí  
User: bueno he estoy encantado de haber hablado contigo y creo que tendrás un gran futuro  
Chatbot: ¡Gracias! ¡Yo también lo espero! User: la verdad adios  
Chatbot: Estaré aquí para ti  
User: adios  
Chatbot: Adiós, lamento las molestias.

## Without RL

User: hola buenas noches  
Chatbot: Buenas noches, dulce príncipe.  
User: cómo te encuentras  
Chatbot: Yo estoy bien gracias por preguntar. ¿Cómo estás?  
User: yo estoy muy bien  
Chatbot: Eso es bueno escuchar. Espero que todo vaya bien con su familia y amigos.  
User: en la familia todos estamos bien y me alegra de que tú estés contenta y feliz  
Chatbot: Muchas gracias. Espero que las cosas también te vayan bien.  
User: quería saber si te gusta mucho el cine  
Chatbot: Ver películas es uno de mis pasatiempos favoritos, ¡así que sí! Las películas siempre son divertidas cuando son buenas películas.  
User: y qué tipo de películas te gustan  
Chatbot: Mayormente comedias y dramas mayormente. A veces veo películas de terror o de acción dependiendo de mi estado de ánimo, etc. Pero sobre todo películas de comedia. Sin embargo, sobre todo comedias.  
User: a mí también me gustan mucho las comedias y además me gusta mucho la música de estas películas  
Chatbot: Sí, también soy un gran fanático de las comedias. Especialmente comedias con temas pesados como Padre de familia y Amigos. La comedia es mi género favorito.  
User: me alegro de que coincidamos y además te gusta el teatro  
Chatbot: ¡Sí! Deberíamos salir pronto. ¿Quizás en algún momento durante las vacaciones de primavera o de verano? O tal vez en cualquier momento después de que terminen las vacaciones escolares.  
Lo que sea que funcione para tí.  
User: y te gusta viajar por los países del mundo  
Chatbot: Sí, he estado viajando bastante últimamente en realidad. Pero no estoy seguro de si vale la pena ya que me he metido en otras cosas además de las películas.  
User: a qué países del mundo te gusta viajar más  
Chatbot: Italia Italia Italia Italia Italia Italia Italia ItalianItalianItaly... [cut sequence]  
User: a mí también me encanta italia sobre todo florencia  
Chatbot: Italia es hermosa. Estoy celoso de la belleza de tu país.  
User: te gustaría ir a venecia conmigo  
Chatbot: Venecia es hermosa, pero Venecia es un poco más cara que Florencia. Venecia es más barata que Florencia. Definitivamente vale la pena visitar Venecia si alguna vez estás en la zona.  
User: y te gustaría venir a roma  
Chatbot: ¡Cosa segura! Roma es hermosa y me encantaría visitar Roma algún día. Roma también es hermosa y me gustaría visitar Roma pronto también.  
User: no pases  
Chatbot: Bien, Roma es hermosa y Roma es hermosa. Roma también es preciosa y Roma también es preciosa. Roma también es bonita y Roma también es preciosa. Roma es preciosa y Roma es maravillosa. Roma es preciosa.  
User: pues prepárate porque vamos a hacer un viaje muy pronto dentro de un mes a roma hola  
Chatbot: ¡Ya estoy allí! Y ya estoy planeando un viaje a Roma el año que viene cuando visite Italia por primera vez. ¡Va a ser increíble!  
User: y no te gustaría ir conmigo yo ya he estado en roma y la conozco muy bien te puedo hacer de cicerone  
Chatbot: ¡Oh, absolutamente! ¡Eso suena maravilloso! :RE  
User: pues quedamos para el mes de agosto del año que viene para ir juntos  
Chatbot: ¡Eso suena genial! ¡Nos veremos entonces!  
User: te advierto que en el mes de agosto hace mucho calor en roma pero  
Chatbot: Sí, estoy seguro de que estará bien. De todos modos, aquí es verano, así que no debería ser tan malo de ninguna manera.  
User: eso te quería decir que hay muchos puestos de helados para refrescarse  
Chatbot: Sí, eso significa. Y las heladerías siempre están abiertas durante el verano porque son más baratas que los restaurantes y cafés. Así que sí, es un gran momento para las heladerías.  
User: y no te gustaría venir conmigo a nueva york  
Chatbot: ¡La ciudad de Nueva York es increíble! Especialmente Manhattan Beach. Pero la ciudad de Nueva York en sí tampoco es sorprendente. Es solo una gran ciudad llena de gente que ama la comida y la bebida.  
User: eso es verdad lo que dices pero también la música y el cine y las comedias musicales y woody allen que te parece  
Chatbot: Woodley Allen es el mejor baterista de todos los tiempos. Periódicamente se mete en discusiones con su propio compañero de batería y ambos se llevan muy bien juntos.  
User: bueno pues quedamos para ir a nueva york dentro de 2 años  
Chatbot: ¡Eso es genial, estoy seguro de que será una maravilla! ¡Diviértete hombre!  
User: bueno y cambiando de tema te gusta el fútbol  
Chatbot: El fútbol es genial, pero el fútbol ya no es lo mío. Sin embargo, el baloncesto es donde está. Y el baloncesto también es bastante tonto.  
User: veo que el deporte no te gusta tanto como el cine  
Chatbot: Sí, tampoco me gustan los deportes. Sin embargo, el baloncesto es definitivamente más popular que el fútbol hoy en día. El baloncesto ha crecido de manera constante durante los últimos años en Estados Unidos. El baloncesto está creciendo exponencialmente en todo el mundo.  
User: estamos de acuerdo bueno estoy encantado de haber hablado todo este rato contigo  
Chatbot: Me alegro de que pudiéramos reunirnos y discutir temas de baloncesto entre nosotros, siempre es bueno cuando la gente puede reunirse en algo relacionado con el baloncesto. ¡Salud!  
User: me ha gustado mucho hablar contigo nos veremos pronto adios  
Chatbot: Saludos amigo, espero que disfrutes el juego tanto como yo. Saludos por charlar conmigo también. ¡Salud!  
User: adios  
Chatbot: ¡Salud! ¡Espero que hayas disfrutado charlando con los dos! ¡Salud!

## With RL (episode 60)

## Dialogues with a real participant

# Human-interaction experiments | Hypothesis

## RL hypothesis

**Null hypothesis  $H_0$**

RL **does not improve** the chatbot engagement

**Alternative hypothesis  $H_A$**

RL **improves** the chatbot engagement

## Age-range hypothesis

**Null hypothesis  $H_0$**

The engagement **is dependent** on the age-range.

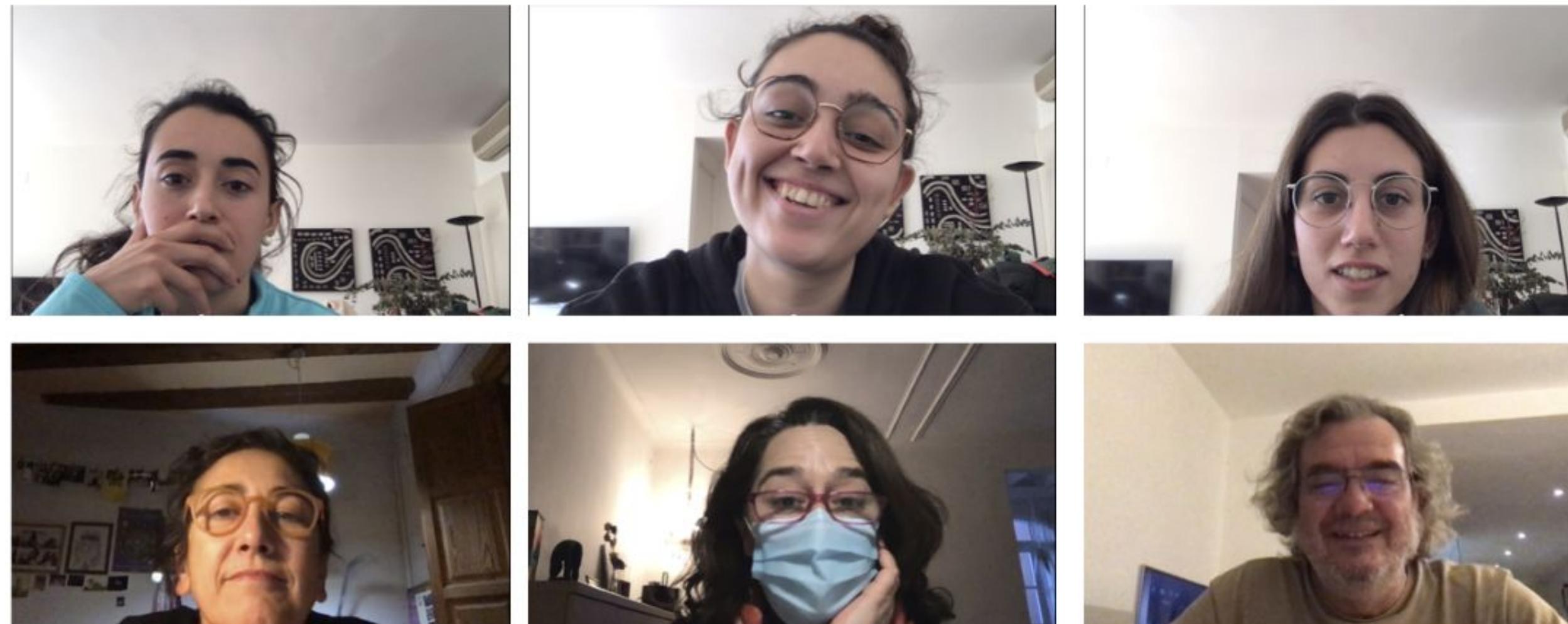
**Alternative hypothesis  $H_A$**

The engagement **does not depend** on the age-range.

# Human-interaction experiments | Participants

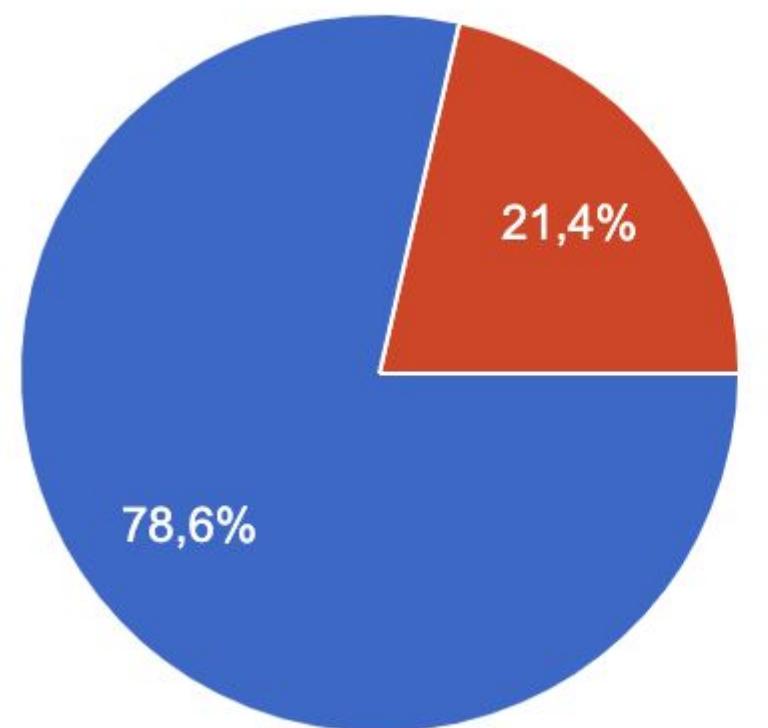
- **14 Spanish subjects (57.1% male /42.9% female)**

- 7 between 20-30 years old
- 7 above 55 years old

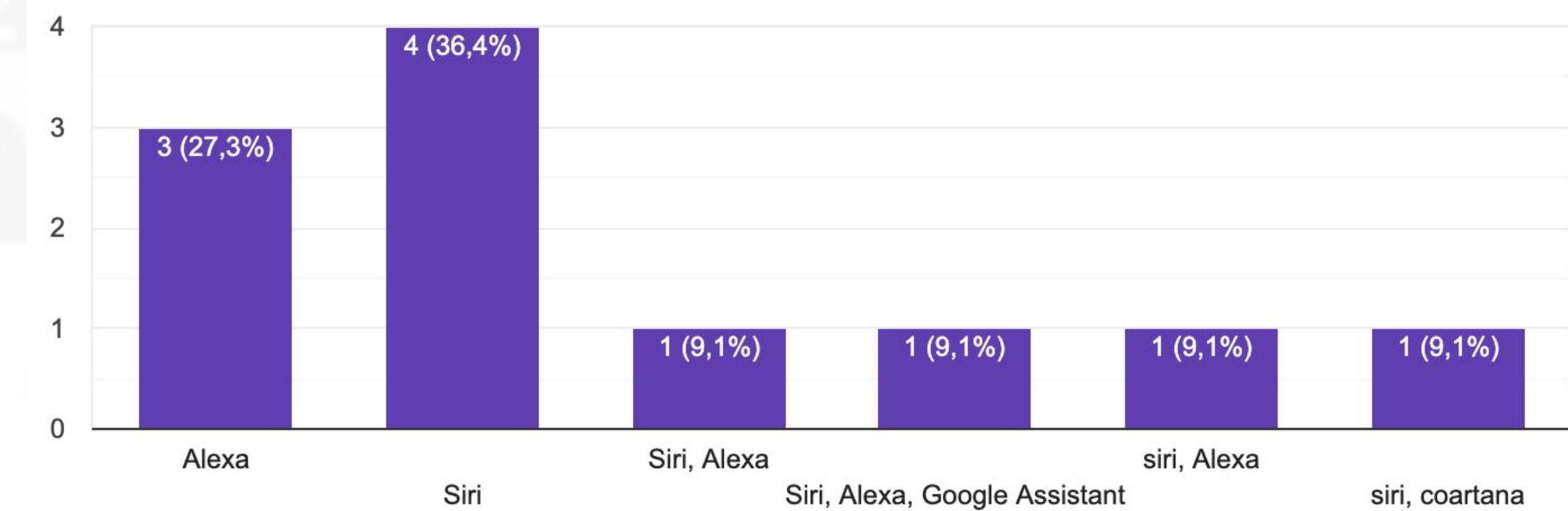


# Human-interaction experiments | Participants

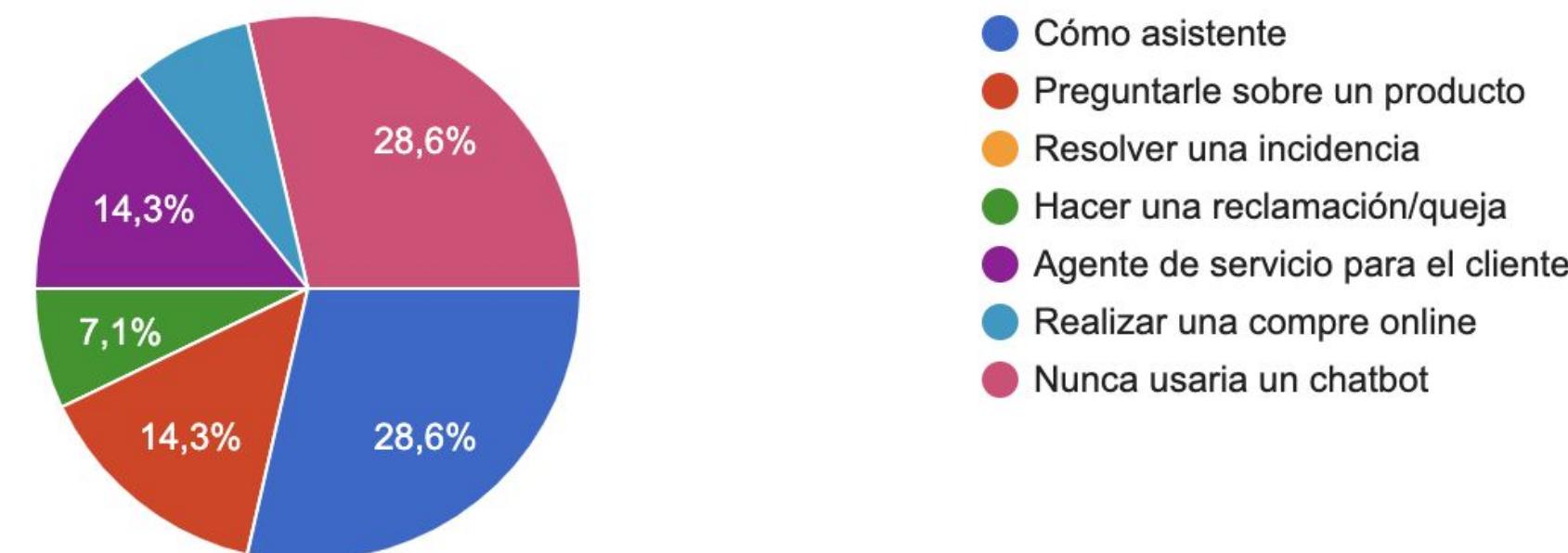
Before this test...Have you ever chatted with a chatbot?



If the previous question was "YES", what chatbots have you used?



What do you think you would use a chatbot for?



# Human-interaction experiments | Study design

## Baseline Model



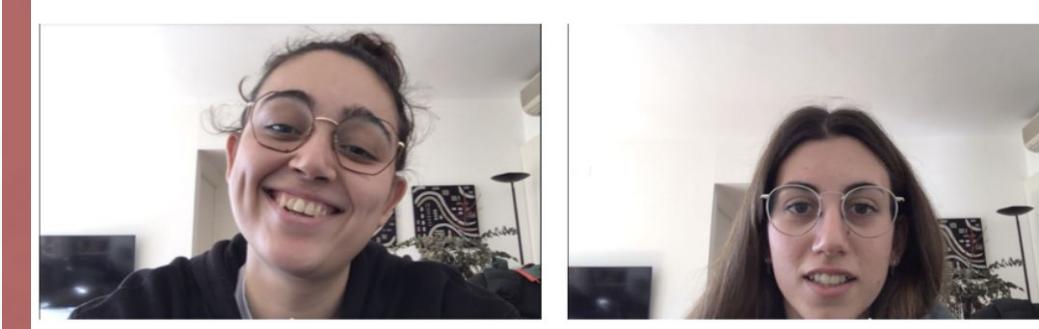
1. Task instructions
2. Record the session
3. Interaction with **simple model**
4. Questionnaire

## RL Model



1. Task instructions
2. Record the session
3. Interaction with the **simple model**
4. Questionnaire

## 20-30 years old



1. Task instructions
2. Record the session
3. Chatbot interaction with the RL model
4. Questionnaire

## +55 years old



1. Task instructions
2. Record the session
3. Chatbot interaction with the RL model
4. Questionnaire

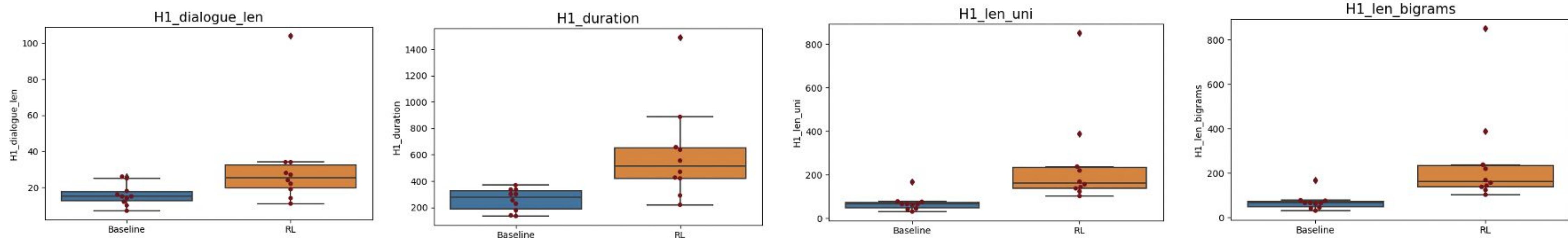


# Human-interaction experiments | Evaluation

## Hypothesis 1: Quantitative evaluation

- RL model gets **better engagement** and **expressiveness** metrics than the baseline model with real users:

Model	Avg. turns	Avg. duration (s)	Avg. unigrams	Avg. bigrams
DialoGPT standard	15.8	257.693	69.7	253.3
DialoGPT with RL	<b>31.7</b>	<b>605.251</b>	<b>253.3</b>	<b>337.5</b>

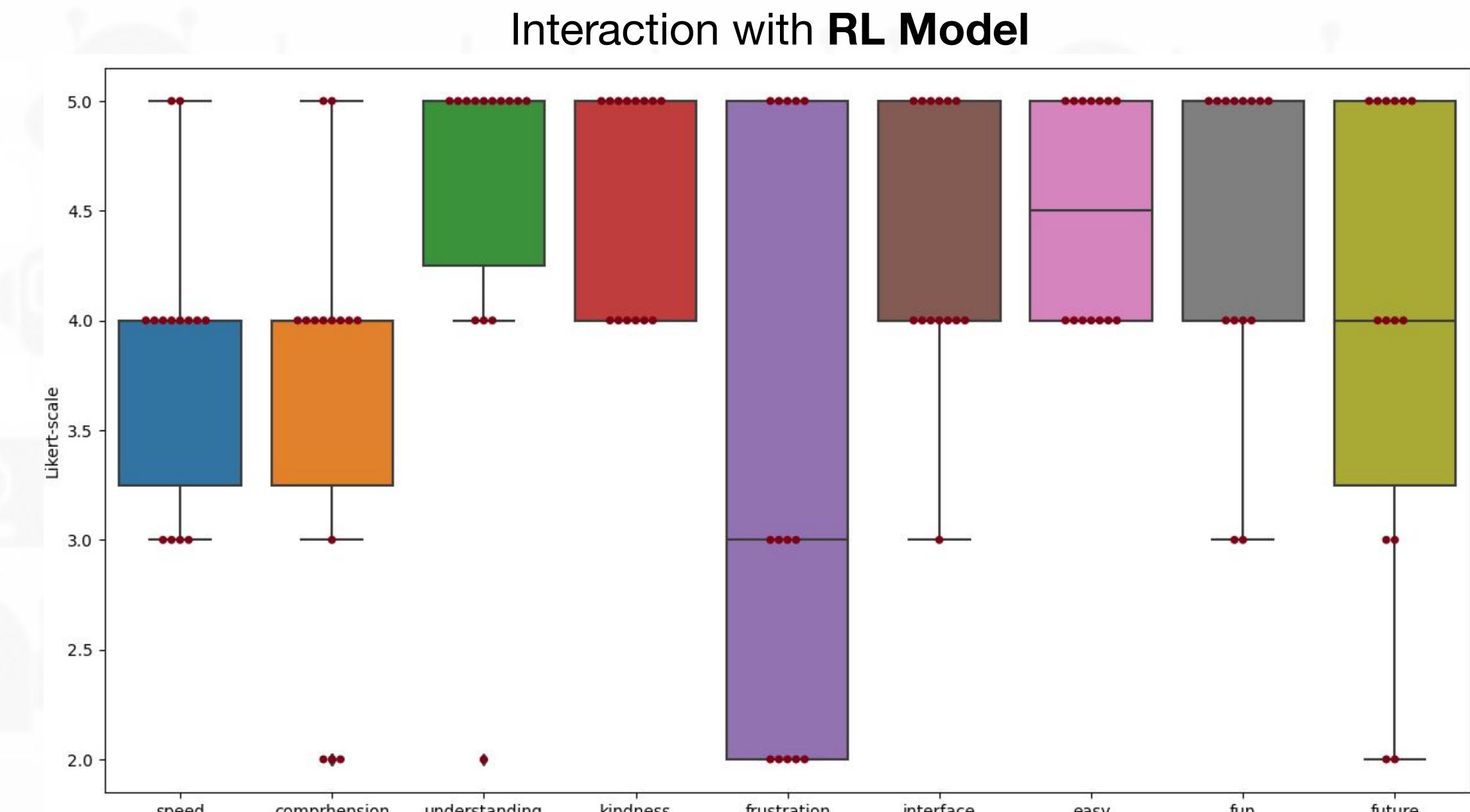
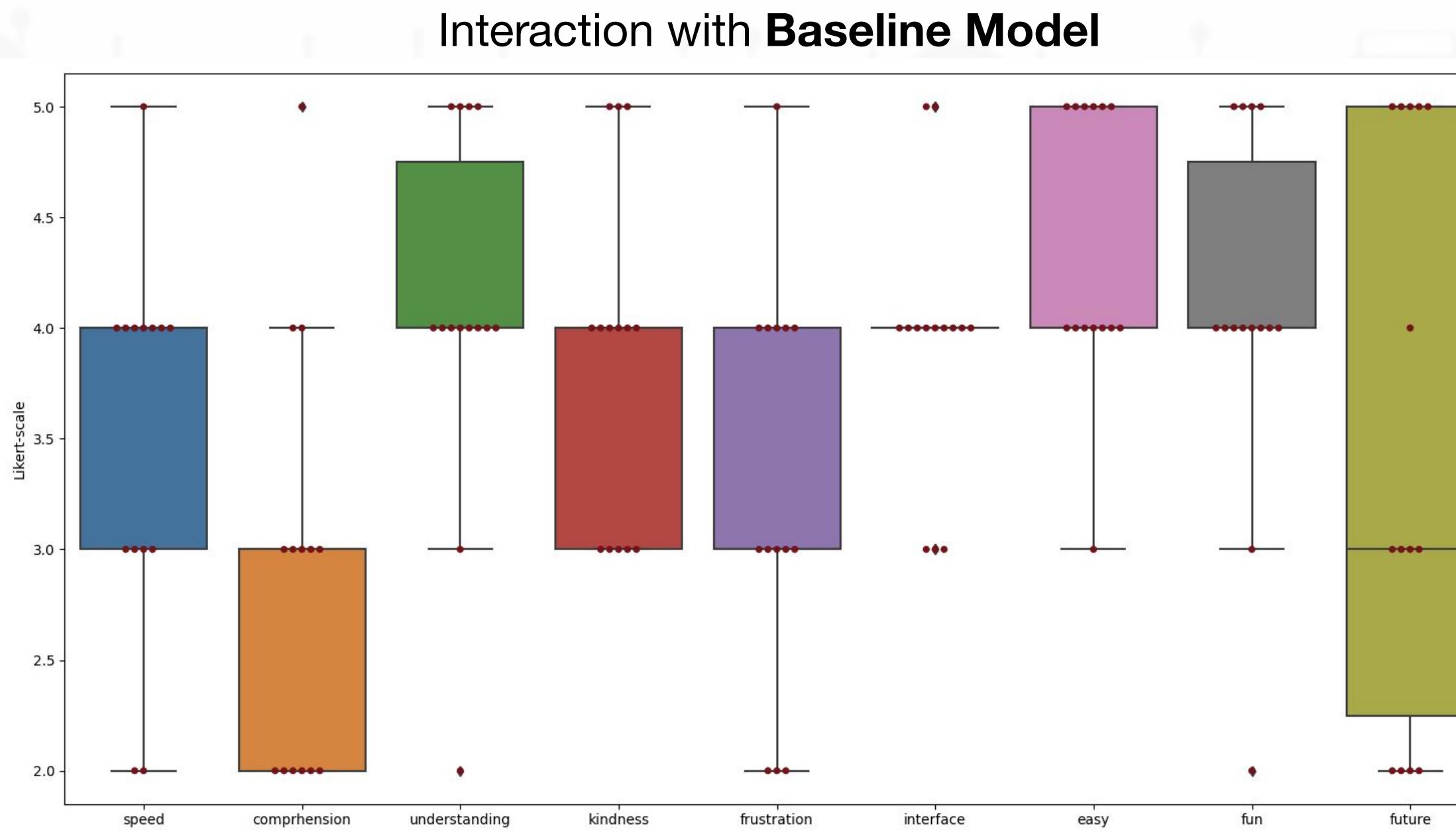


# Human-interaction experiments | Evaluation

## Hypothesis 1: Qualitative evaluation

Results from **likert-scale questions** about chatbot engagement attributes:

- Speed, comprehension, understanding, kindness, frustration, interface, easiness, fun, future

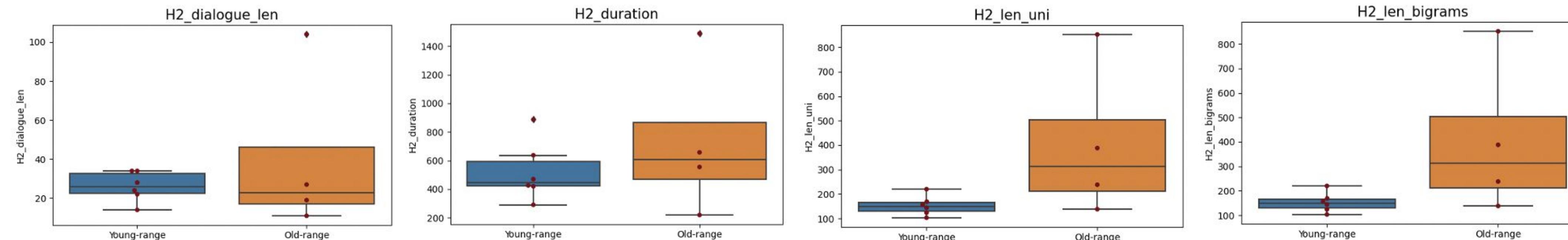


# Human-interaction experiments | Evaluation

## Hypothesis 2: Quantitative evaluation

- Old people show significantly **more engagement** than young participants

Age-range	Avg. turns	Avg. duration (s)	Avg. unigrams	Avg. bigrams
Young (20-30)	26.0	521.972	152.833	243.166
Old (+55)	<b>40.25</b>	<b>730.170</b>	<b>404.0</b>	<b>479.0</b>

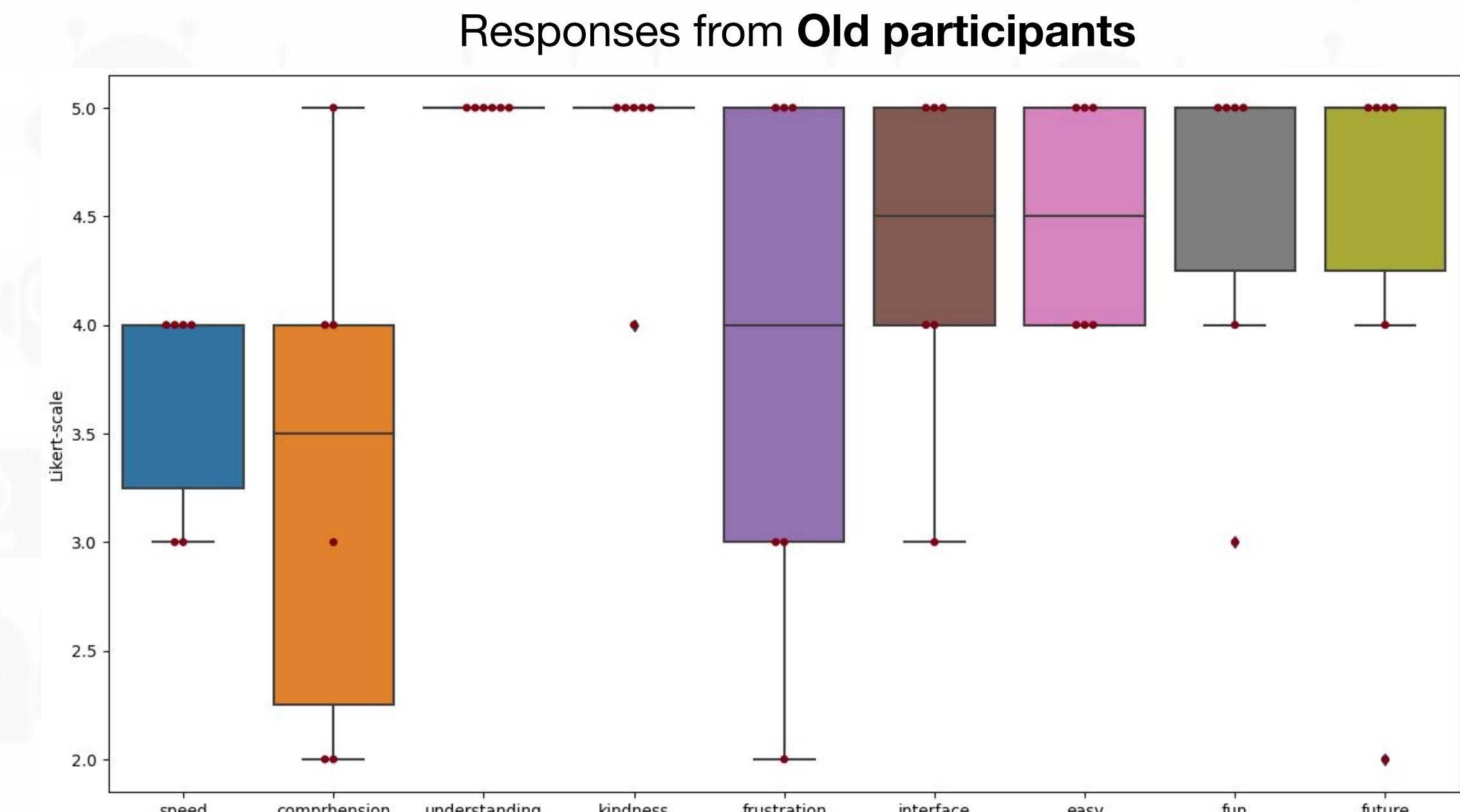
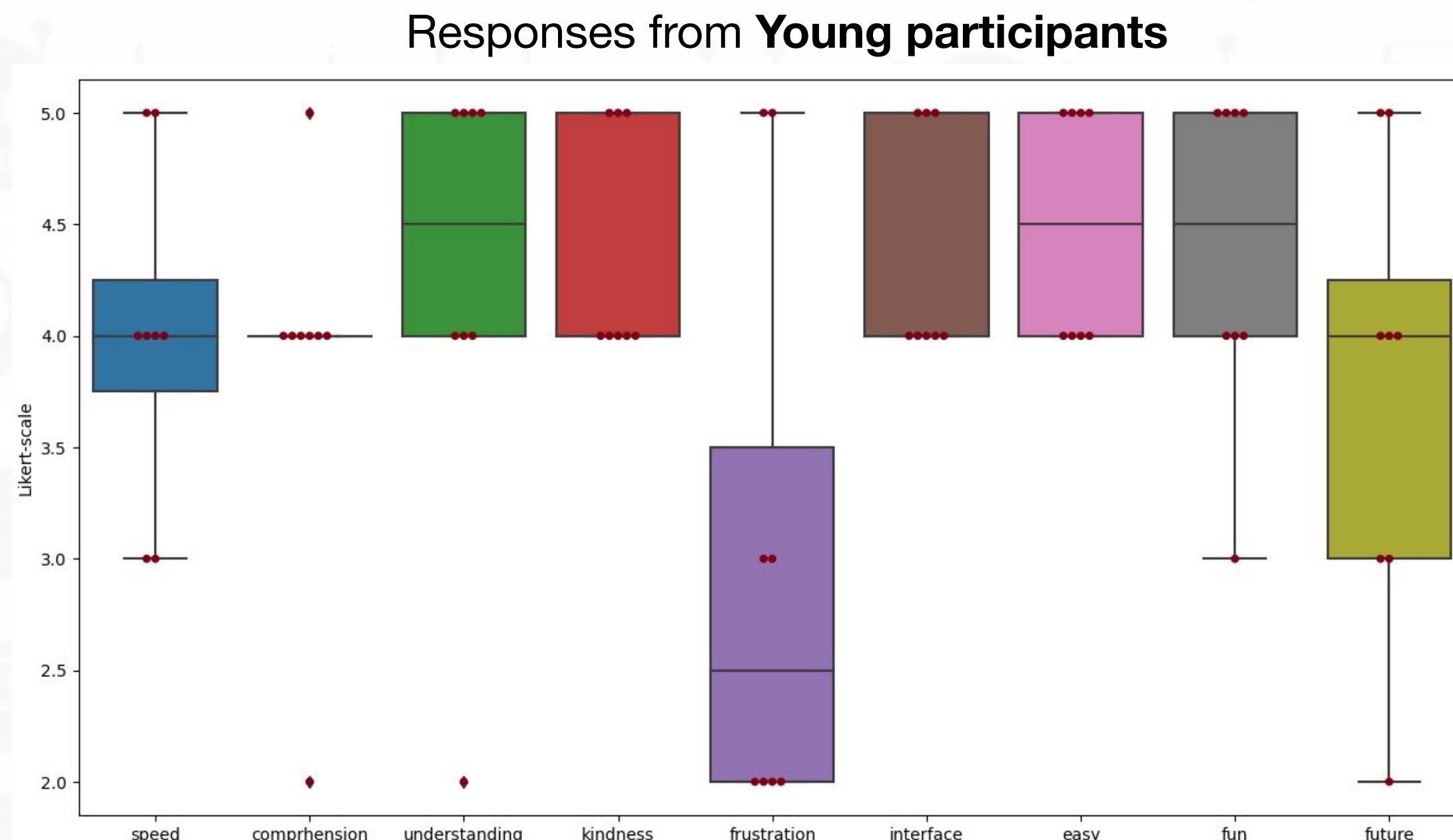


# Human-interaction experiments | Evaluation

## Hypothesis 2: Qualitative evaluation

Results from **likert-scale questions** about chatbot engagement attributes:

- Speed, comprehension, understanding, kindness, frustration, interface, easiness, fun, future



# Human-interaction experiments | Discussion

- We didn't have enough experiments to statistically reject the hypotheses
- ( $H_1$ ) With **RL**, the model is able to perform **more expressive** and **longer sentences**, leading to **more engagement** during the conversation
  - More dialogue turns and duration
  - More word unigrams and bigrams
  - Likert-scale questions give higher scores (except *frustration*)
- ( $H_2$ ) Interaction with **elder people** appear to show **more engagement** than younger people
  - More dialogue turns and duration
  - More word unigrams and bigrams
  - Likert-scale questions give higher scores (except *comprehension & frustration*)



# Conclusion & Future work |

- We managed to successfully **implement, build and test** our **open-domain voice-based chatbot** prototype.
- We have demonstrated that the model benefits from the RL training and **increases the engagement** among the study subjects
- We have seen that the model is able to respond to **many different topics** (movies, music, sports, countries, etc), proving effective response to **open-domain**.
- Because we did our experiments with Spanish-native subjects, **utterance translation** was not always accurate
- **Future work**
  - Improve speech recognition (strange loading behaviour)
  - Improve translations from Spanish to English to use the model in other languages
  - Try larger Transformers (DialogPT-Large)

# References

- ZHANG, Yizhe, et al. DialoGPT: Large-scale generative pre-training for conversational response generation. *arXiv preprint arXiv:1911.00536*, 2019.
- ADIWARDANA, Daniel, et al. Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*, 2020.
- SERBAN, Iulian V., et al. A deep reinforcement learning chatbot. *arXiv preprint arXiv:1709.02349*, 2017.
- ZHOU, Li, et al. The design and implementation of xiaoice, an empathetic social chatbot. *Computational Linguistics*, 2020, vol. 46, no 1, p. 53-93.
- YU, Dian, et al. Gunrock: A social bot for complex and engaging long conversations. *arXiv preprint arXiv:1910.03042*, 2019
- SHANG, Lifeng; LU, Zhengdong; LI, Hang. Neural responding machine for short-text conversation. *arXiv preprint arXiv:1503.02364*, 2015.
- SORDONI, Alessandro, et al. A neural network approach to context-sensitive generation of conversational responses. *arXiv preprint arXiv:1506.06714*, 2015.
- VINYALS, Oriol; LE, Quoc. A neural conversational model. *arXiv preprint arXiv:1506.05869*, 2015.
- SERBAN, Iulian V., et al. Building end-to-end dialogue systems using generative hierarchical neural network models. *arXiv preprint arXiv:1507.04808*, 2015.
- ZHANG, Yizhe, et al. Dialogpt: Large-scale generative pre-training for conversational response generation. *arXiv preprint arXiv:1911.00536*, 2019.
- SO, David R.; LIANG, Chen; LE, Quoc V. The evolved transformer. *arXiv preprint arXiv:1901.11117*, 2019.
- LI, Jiwei, et al. Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*, 2016.

# Human-interaction experiments | Demo

## IN VIVO DEMONSTRATION

# Solution proposal | Cognitive frame

## PROBLEM

Neural models of dialogue generation usually tend to be short-sighned, predicting utterances one at time getting stuck in an infinite loop of repetitive responses, therefore, they ignore the influence of future outcomes.

### STEP 01

Use a large **neural conversational response generation model**.

DialoGPT (Dialogue Generative pre-trained transformer)

### STEP 02

Use deep reinforcement learning to **model future reward** in chatbot dialogue.

### STEP 03

**Train the DialoGPT model by dialoguing to improve long-term reward with a simulated agent (DialoGPT).**