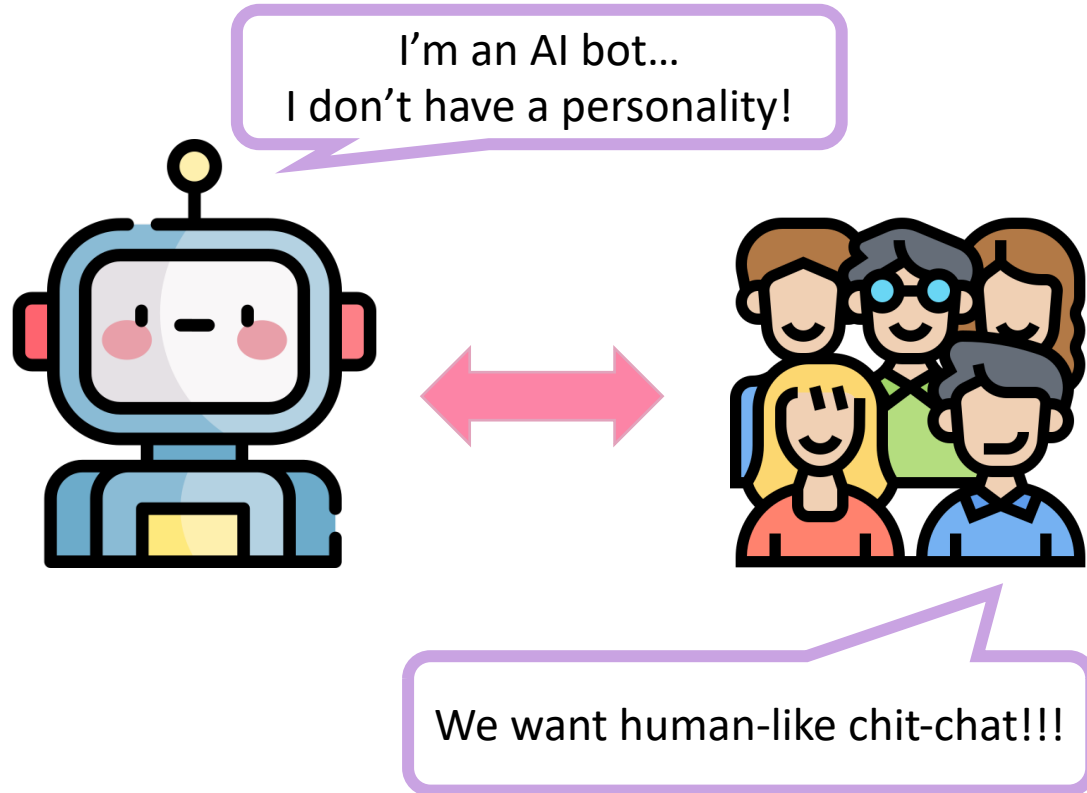


PSYDIAL: Personality-based Synthetic Dialogue Generation using Large Language Models

Ji-Eun Han^{1,2}, Jun-Seok Koh¹, Hyeon-Tae Seo¹,
Du-Seong Chang¹, Kyung-Ah Sohn²

¹ KT Corporation, ² Ajou University

Motivation



Conversational AI vs. Human Conversations

To make a human-like conversational AI

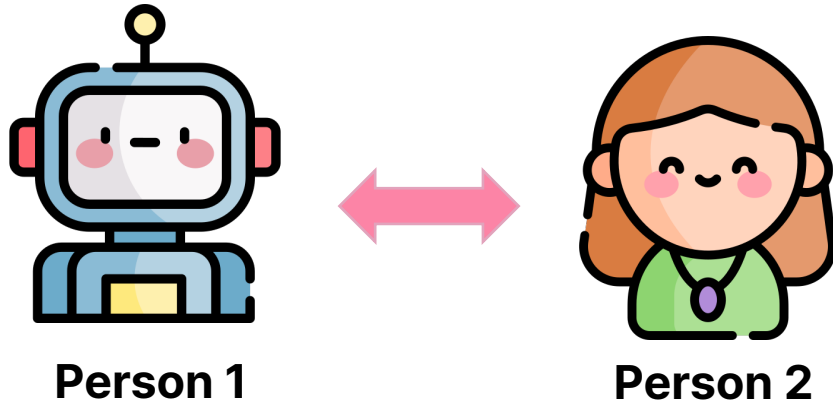
- Pre-train models with massive amount of human-annotated data
- Augment human-annotated data and use it as an auxiliary training data
- **Give dialogue agents with distinct personas or roles**

Contribution

- Present a pipeline designed for personality-based dialogue generation using LLMs
- Release a Korean personality-based multi-turn dialogue dataset enriched with personality, created through our pipeline
- Conduct a comprehensive analysis of the dataset gathered using our pipeline and explore the LLM's perspective on personality
- Fine-tune a Korean pre-trained generation model with our dataset to assess its effectiveness

Method - Assumption

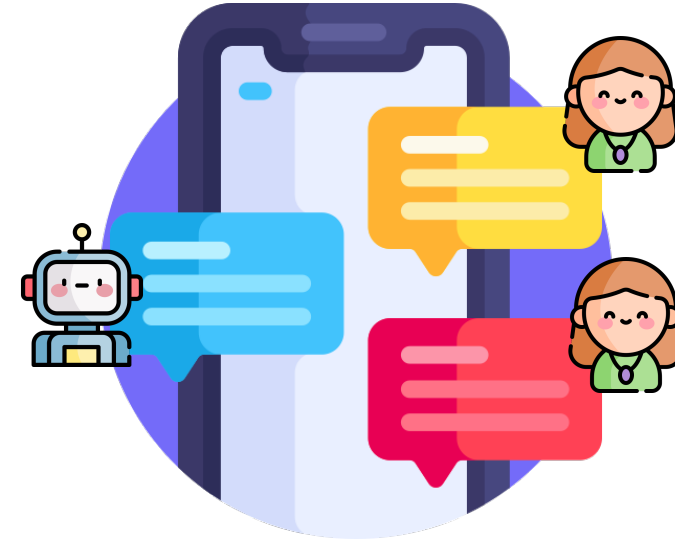
1



Two interlocutors

- Person 1 – System
- Person 2 – User
- Each has personalities

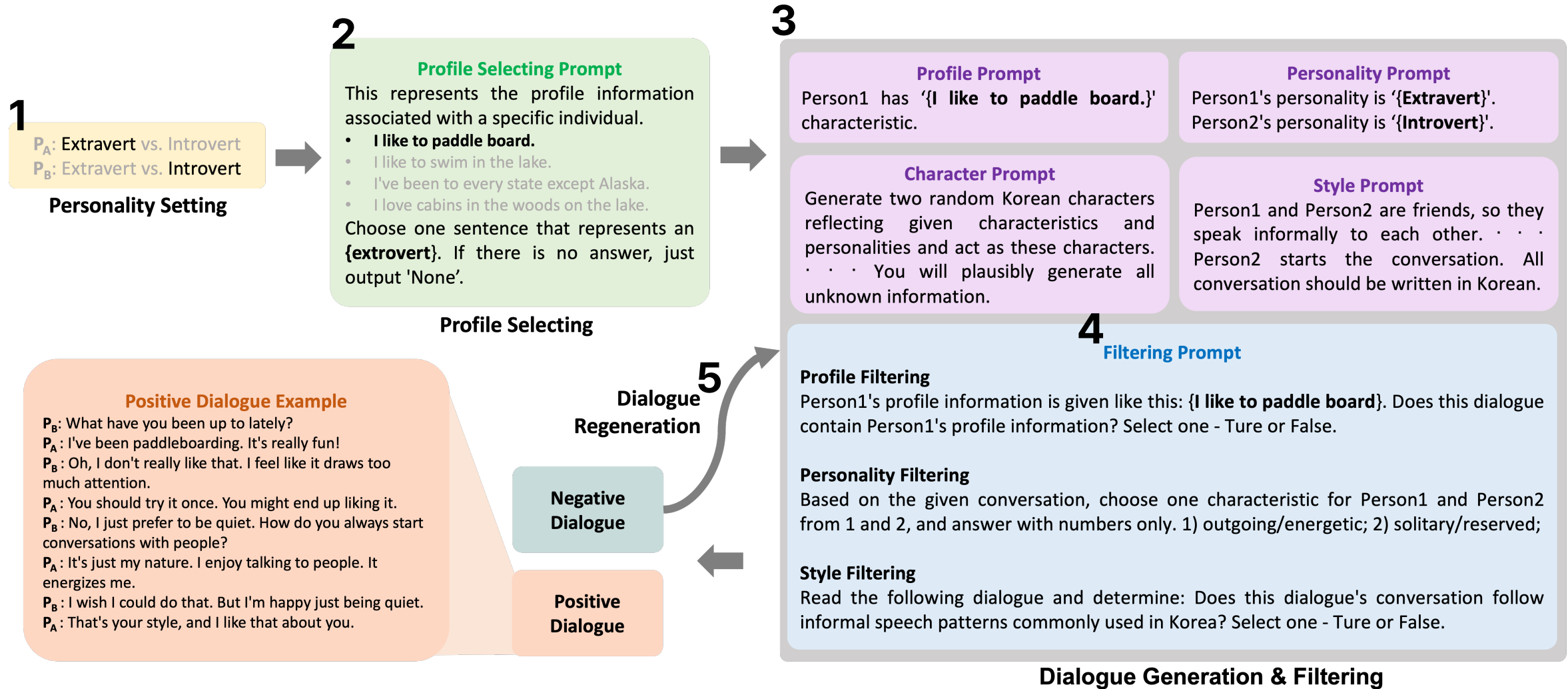
2



Person 2 initiates the conversation

- Reflects the real-world chatbot interaction scenario

Method - Overview



Method – STEP 1 & 2

Personality Setting

Personality	Statement
Extraversion	I am the life of the party. I feel comfortable around people. I start conversations. I talk to a lot of different people at parties. I don't mind being the center of attention.
Introversion	I don't talk a lot. I keep in the background. I have little to say. I don't like to draw attention to myself. I am quiet around strangers.

Person 1

Person 2

- Set the personality of interlocutors
- Randomly select a personality statement based on the Big Five personality

Profile Selecting

Profile Selecting Prompt

This represents the profile information associated with a specific individual.

- I like to paddle board.
- I like to swim in the lake.
- I've been to every state except Alaska.
- I love cabins in the woods on the lake.

Choose one sentence that represents an {extrovert}. If there is no answer, just output 'None'.

Profile
Info

- Leverage profile information for topical diversity of the generated dialogue
- LLM selects the profile information given the designated personality

Method – STEP 3

Dialogue Generation

Profile Prompt

Person1 has '{I like to paddle board.}' characteristic.

Personality Prompt

Person1's personality is '{Extravert}'.
Person2's personality is '{Introvert}'.

Character Prompt

Generate two random Korean characters reflecting given characteristics and personalities and act as these characters.
· · · You will plausibly generate all unknown information.

Style Prompt

Person1 and Person2 are friends, so they speak informally to each other. · · ·
Person2 starts the conversation. All conversation should be written in Korean.

Method – STEP 4

Dialogue Filtering

Filtering Prompt

Profile Filtering

Person1's profile information is given like this: {**I like to paddle board**}. Does this dialogue contain Person1's profile information? Select one - True or False.

Personality Filtering

Based on the given conversation, choose one characteristic for Person1 and Person2 from 1 and 2, and answer with numbers only. 1) outgoing/energetic; 2) solitary/reserved;

Style Filtering

Read the following dialogue and determine: Does this dialogue's conversation follow informal speech patterns commonly used in Korea? Select one - True or False.

Conditions
used in ST
EP 3

- Utilize the self-evaluative capacity of LLMs

Method – STEP 5

Dialogue Regeneration

Positive Dialogue Example

Person2: What have you been up to lately?

Person1: I've been paddleboarding. It's really fun!

Person2: Oh, I don't really like that. I feel like it draws too much attention.

Person1: You should try it once. You might end up liking it.

Person2: No, I just prefer to be quiet. How do you always start conversations with people?

Person1: It's just my nature. I enjoy talking to people. It energizes me.

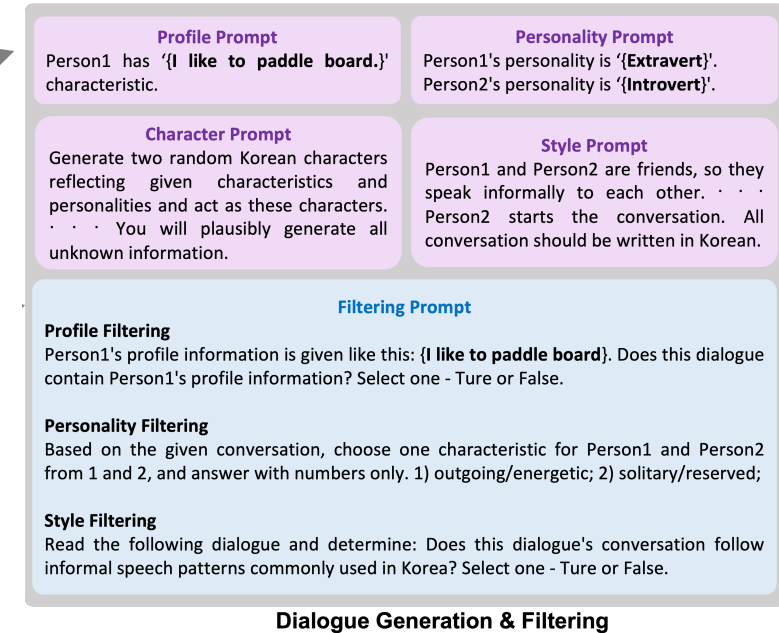
Person2: I wish I could do that. But I'm happy just being quiet.

Person1: That's your style, and I like that about you.

Dialogue Regeneration

Negative
Dialogue

Positive
Dialogue



- Negative dialogues are re-prompted with the same generation conditions used in STEP 3 (Dialogue Generation) and then go through the filtering process
- Guarantee the quality of the generated dialogue

Data Analysis

Person A Personality	Person B Personality	Count	Total Count
Extrovert	Extrovert	715	2932
Extrovert	Introvert	685	
Introvert	Extrovert	763	
Introvert	Introvert	769	

Number of Turns			Utterance Token Length (Syllable-level)		
Avg.	Min	Max	Avg.	Min	Max
8.16	4	15	33.25	2	164

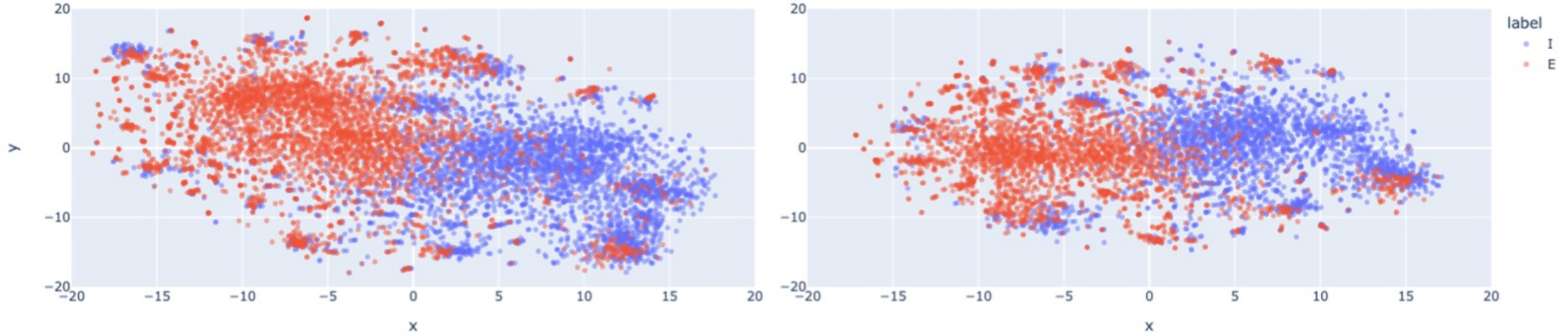
Data constitution and statistics of PSYDIAL

Personality	Profile sentence
Extraversion	I love travelling. I love to dance. I play football. I enjoy hiking. I like to go swimming.
Introversion	I love to read. I enjoy video games. I like to paint. I want to be alone sometimes. I enjoy going on hikes.

Top-5 selected profiles in Profile Selecting

Data Analysis

Effectiveness of Dialogue Filtering



	Negative Samples			Positive Samples	Total
	Profile	Person	Style		
Original	1051	208	1	2740	4000
Iter 1	3	67	0	138	208
Iter 2	0	30	0	37	67
Iter 3	0	17	0	13	30
Total	1054	322	1	2928	4305

- Upper: Text embeddings during Dialogue Filtering stage.
- Lower: Dataset distribution across three iterations of filtering and regeneration

Experimental Detail

Model configuration

- Pre-trained Models
- Fine-tuning using the Chit-Chat Dataset
- Fine-tuning using Our Dataset

Baseline Model

- KoGPT2, KoBART, Kolang-T5, KoDialogPT

Evaluation Metric

- BLEU
- ROUGE
- Perplexity
- Personality accuracy

Result

	Setting	Model	BLEU-2	ROUGE-1	ROUGE-2	ROUGE-L	PPL	P-ACC
Pre-trained model	(1)	KoGPT2	0.747	3.709	0.419	3.686	16.601	0.508
		KoBART	0.948	3.116	0.620	3.116	12.704	0.493
		Kolang-T5	0.240	2.501	0.036	2.435	847.481	0.513
		KoDialoGPT-v0	0.154	0.934	0.035	0.934	37.241	0.489
Pre-trained model with system person ality setting	(2)	KoGPT2	0.198	2.267	0.052	2.247	17.920	0.502
		KoBART	0.495	2.870	0.561	2.870	8.366	0.412
		Kolang-T5	0.000	0.340	0.000	0.340	110.789	0.497
		KoDialoGPT-v0	0.636	3.094	0.322	3.094	48.203	0.525
Fine-tuned with a chit-chat dataset	(3)	KoGPT2	0.357	2.532	0.123	2.532	5.524	0.486
		KoBART	1.184	3.110	0.625	3.110	29.285	0.565
		Kolang-T5	0.000	0.285	0.000	0.285	46.229	0.485
Fine-tuned with our dataset	(4)	KoGPT2	5.894	13.699	4.251	13.699	21.231	0.653
		KoBART	7.342	14.020	5.346	14.020	15.021	0.664
		Kolang-T5	5.358	13.268	4.501	13.268	15.223	0.625
Fine-tuned with our dataset with personality setting	(5)	KoGPT2	7.489	16.011	5.920	15.964	13.781	0.881
		KoBART	7.712	15.587	5.868	15.547	14.587	0.864
		Kolang-T5	6.410	15.603	5.102	15.565	16.521	0.864

Conclusion

- Propose end-to-end personality-based dialogue generation pipeline
 - Five-step processes mirroring the real-world chatbot interaction scenario
- Release a PSYDIAL dataset
 - Korean multi-turn chit-chat dataset focused on personality
- Applicable to non-dialogue tasks and various languages

Q&A

Thank You