Empowering AAC Users: A Systematic Integration of Personal Narratives with Conversational AI

Sayantan Pal[†], Souvik Das[†], Rohini K. Srihari[†], Jeffery Higginbotham[‡], Jenna Bizovi[‡]

Department of Computer Science and Engineering[†], Department of Communicative Disorders and Sciences[‡], State University of New York at Buffalo

Introduction

Research Goal: AAC users often face **slow communication rates**, typically less than **10 words per minute**, which leads to frustration and isolation. Our study aims to bridge this gap by leveraging Conversational AI to generate **personalized responses** on behalf of AAC users, reflecting their **unique experiences** and intent. By integrating **personal narratives** into AI-driven conversation aids, we strive to make communication **faster** and more **meaningful**, ensuring that the **responses are deeply personalized** and **contextually relevant** to each user.

Methodology: Our approach is twofold.

- First, we build a **custom dataset** from the personal narratives of AAC user. This dataset is then used to **fine-tune encoder-decoder models** (Flan-T5) to generate personalized responses.
- Second, we implement a **RAG**-based method to **retrieve context** from the AAC user's **authored content**, ensuring that the generated responses are contextually relevant and reflect the user's personal experiences. This dual approach enables both **personalized** and accurate response generation, addressing the complex communication needs of AAC users.

EXAMPLES - Our model responses are personalized and addresses the needs of the AAC User

EXAMPLE 1

Dialogue History (Prompt):
Partner: What other things did you do as a kid?
User: If we didn't go to the mall, I would stay home and play with my own toys.
Partner: That makes sense. What was that like for you?
User: I had my own space to play in, and I would play for hours and hours.
Partner: Were you alone when you played?
User: Sometimes, but if I didn't want to play alone I would go downstairs and watch my dad play with his trains.
Partner: Play with his trains?

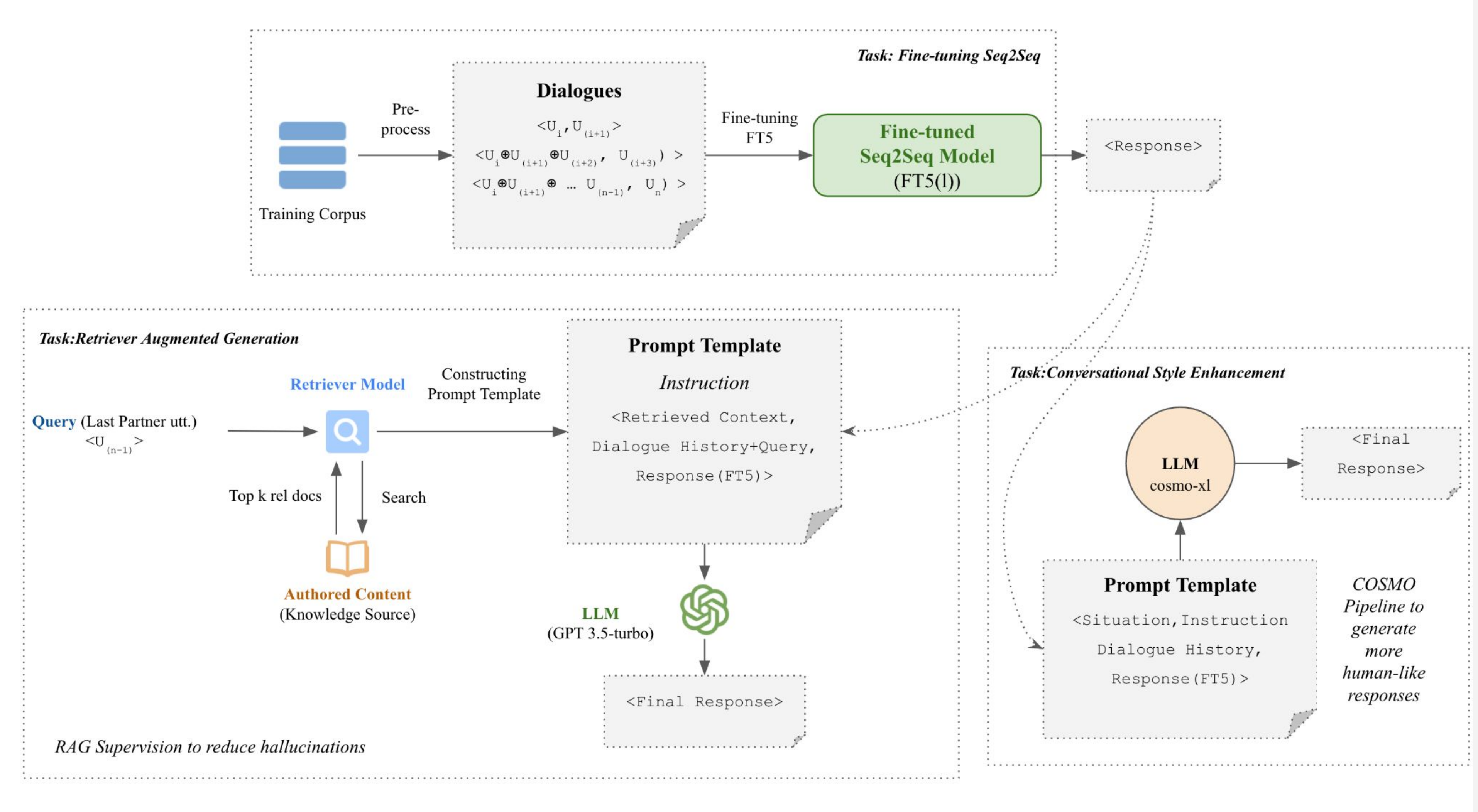
Golden Response: Yes. He has about 200 miniature trains. He would only play with them on weekends.

Baseline Response(Ft. FlanT5): I would watch him play and then we'd go upstairs to our room.

Our Best Model Response (RAG): I'd watch him play with his **200 miniature trains** and then we'd go upstairs and play pretend games.

Note: We can observe that our best model can capture most of the information accurately and satisfy the AAC User.

Architecture



Overview of the Model Architectures. The block at the top shows the fine-tuned Flan-T5(FT5) model on the conversation Dataset. We use the generated response in the next two tasks to perform RAG and conversation style enhancement. Please note here, u means utterance.

Evaluation & Conclusion

Evaluation:
We employed a two-fold evaluation combining **automatic metrics** (e.g., BLEU, METEOR, BERTScore, etc) to gauge text similarity and **human evaluations** based on relevance, sincerity,realism, etc rated by AAC users and judges.

Conclusion:
Our approach demonstrates that integrating personalized AI with AAC systems can significantly enhance communication, aligning generated responses with individual user narratives. This personalized model sets a foundation for extending AI-driven aids to meet diverse communication needs.

Limitations:

- The study's focus on a **single AAC user** limits its generalizability.
- Future work will explore dataset diversity and modular customization to broaden applicability across AAC users, enhancing adaptability.

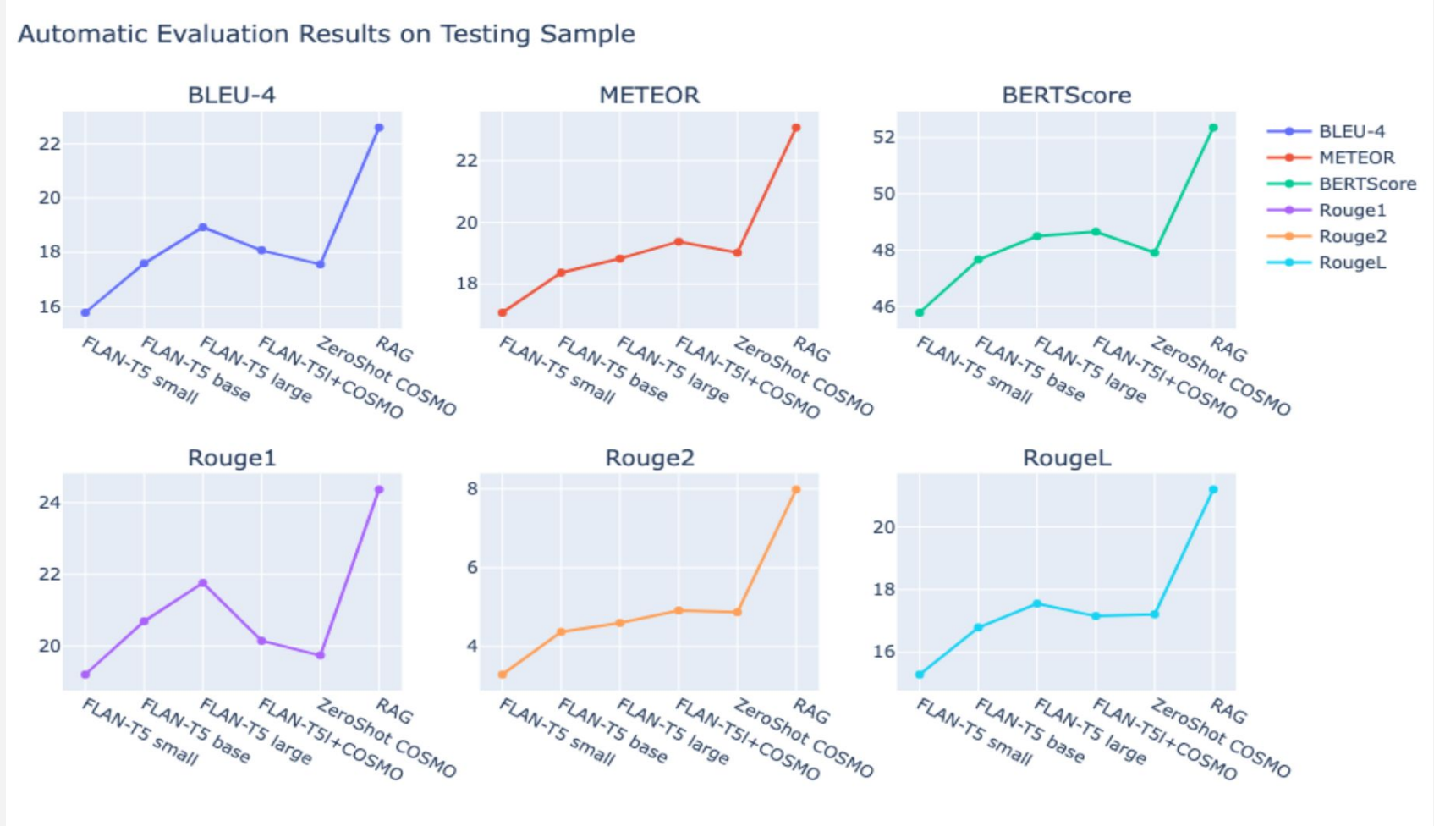
Results

Model	BLEU-4	METEOR	BERTScore	Rouge1	Rouge2	RougeL	Avg.
FT5-zero-shot (s)	10.98	10.12	42.83	12.60	02.67	11.27	15.07
FT5-fine-tuned (s)	15.78	17.08	45.78	19.21	03.29	15.29	19.40
FT5-zero-shot (b)	09.07	10.27	43.43	13.08	03.28	12.06	15.19
FT5-fine-tuned (b)	17.59	18.37	47.66	20.69	04.37	16.79	20.91
FT5-zero-shot (l)	08.09	11.06	44.11	15.13	04.25	14.28	16.15
FT5-fine-tuned (l)	18.93	18.83	48.49	21.76	04.60	17.55	21.69
FT5(l)+cosmo-xl	18.07	19.38	48.65	20.15	04.91	17.15	21.45
cosmo-xl(zero-shot)	17.56	19.02	47.91	19.74	04.87	17.21	21.05
RAG(Llama2-13B)	15.91	17.79	47.84	19.09	05.76	16.54	20.48
RAG(FT5(l)+GPT3.5t)	22.61	23.08	52.36	24.37	07.99	21.20	**25.26**

Automatic Evaluation Results on the Testing Sample (400 Prompt-Response Pairs). FT5 is the Flan-T5 model, and s, b, and l denote small, base, and large configurations, respectively.

Model	Specific	Sincere	Understandable	Relevant	Fluency	Quantity
FT5 (l) J1	1.130	0.900	0.730	0.970	0.800	1.070
FT5 (l) J2	1.110	0.930	0.770	1.000	0.830	1.030
FT5 (l) Avg.	1.120	0.915	0.750	0.985	0.815	1.050
FT5(l)+COSMO J1	0.930	0.930	0.870	1.000	1.030	0.970
FT5(l)+COSMO J2	0.930	0.930	0.870	1.000	1.030	0.970
FT5(l)+COSMO Avg.	0.930	0.930	0.870	1.000	1.030	0.970
COSMO(zero-shot) J1	0.930	0.830	0.870	0.830	1.100	1.000
COSMO(zero-shot) J2	1.030	0.870	0.870	0.830	1.100	0.870
COSMO(zero-shot) Avg.	0.98	0.850	0.870	0.830	1.100	0.935
RAG(FT5(l)+GPT3.5t) J1	1.300	1.230	0.970	1.300	1.400	1.000
RAG(FT5(l)+GPT3.5t) J2	1.300	0.900	0.900	1.300	1.370	1.000
RAG(FT5(l)+GPT3.5t) Avg.	**1.300**	**1.050**	**0.935**	**1.300**	**1.385**	**1.000**

The average of each criterion from the respective judges (30 responses)



Automatic Evaluation Results on Testing Samples compared across different Models.

Average of each criterion from the Respective Judges compared across four different models.

Criterion	Score
Relevant	3.30
Factual	3.40
Concise	3.40
Representative	3.00
Realistic	3.56

The average of each criterion from the AAC User on a 5 point scale on the best model. Where 5 means highest, 1 means lowest(30 responses)

Prompt	Golden	Ft.FT5 (l)	Ft.FT5(l) + COSMO	RAG
Partner: I suppose with PET, you never quite feel alone. Does she follow you? User: As I'm in the shower, I peek out and see PET curled up with the floor mat. She notices me and jumps on the side of the bathtub, pacing back and forth. Partner: Is PET approachable or does she keep her distance? Give an example of your interaction.	User: Turning my head, I meet PET's gaze, and to my surprise, she plants a tiny kiss on the front of my nose.	User: I try to get her attention, but she quickly turns away.	User: She does tend to keep her distance at times, but she's always happy to see me when I come home. She'll start talking to me on the phone even if I don't respond.	User: She gives me a little kiss on the nose, then waits for me.

Conversation example (Any recognizable entities have been anonymized using uppercase characters for the specific entity)

EMNLP 2024 Workshop CustomNLP4U - Paper 10