UNIVERSITY OF LEEDS

UK Research and Innovation

Score Before You Speak:
Improving Persona Consistency in Dialogue Generation using Response Quality Scores

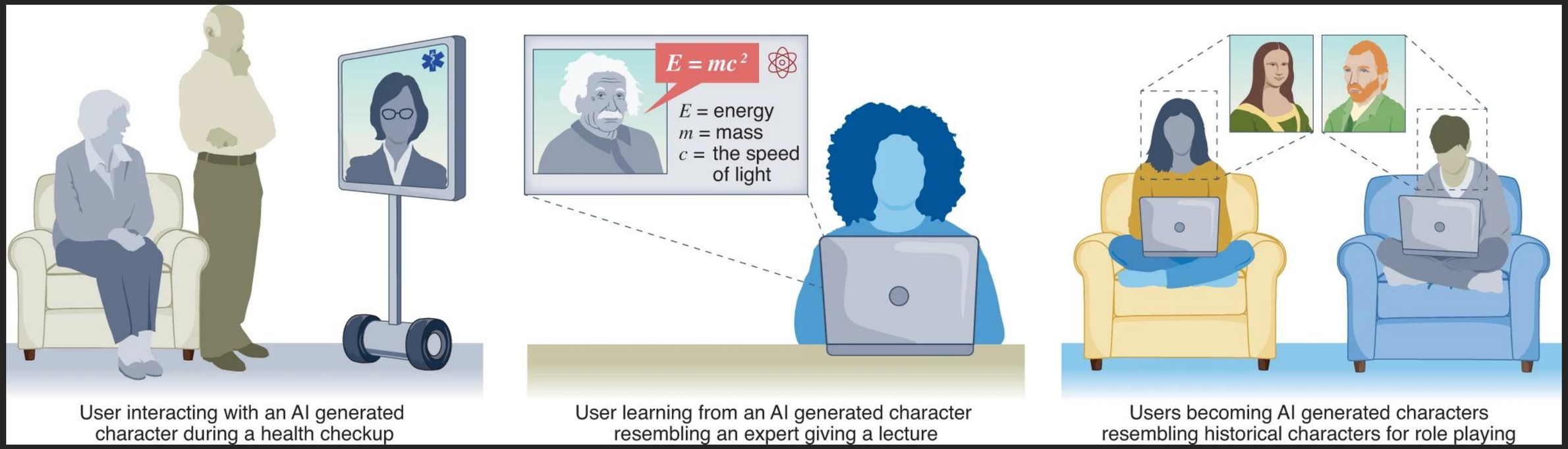
Arpita Saggar[1], Jonathan C. Darling[2], Vania Dimitrova[1],
Duygu Sarikaya[1] and David C. Hogg[1]

[1]School of Computer Science, University of Leeds
[2]Leeds Institute of Medical Education, School of Medicine, University of Leeds

Persona-Based LLM Agents

Generate responses consistent with an assigned persona

Persona-Based LLM Agents: Applications



User interacting with an AI generated character during a health checkup

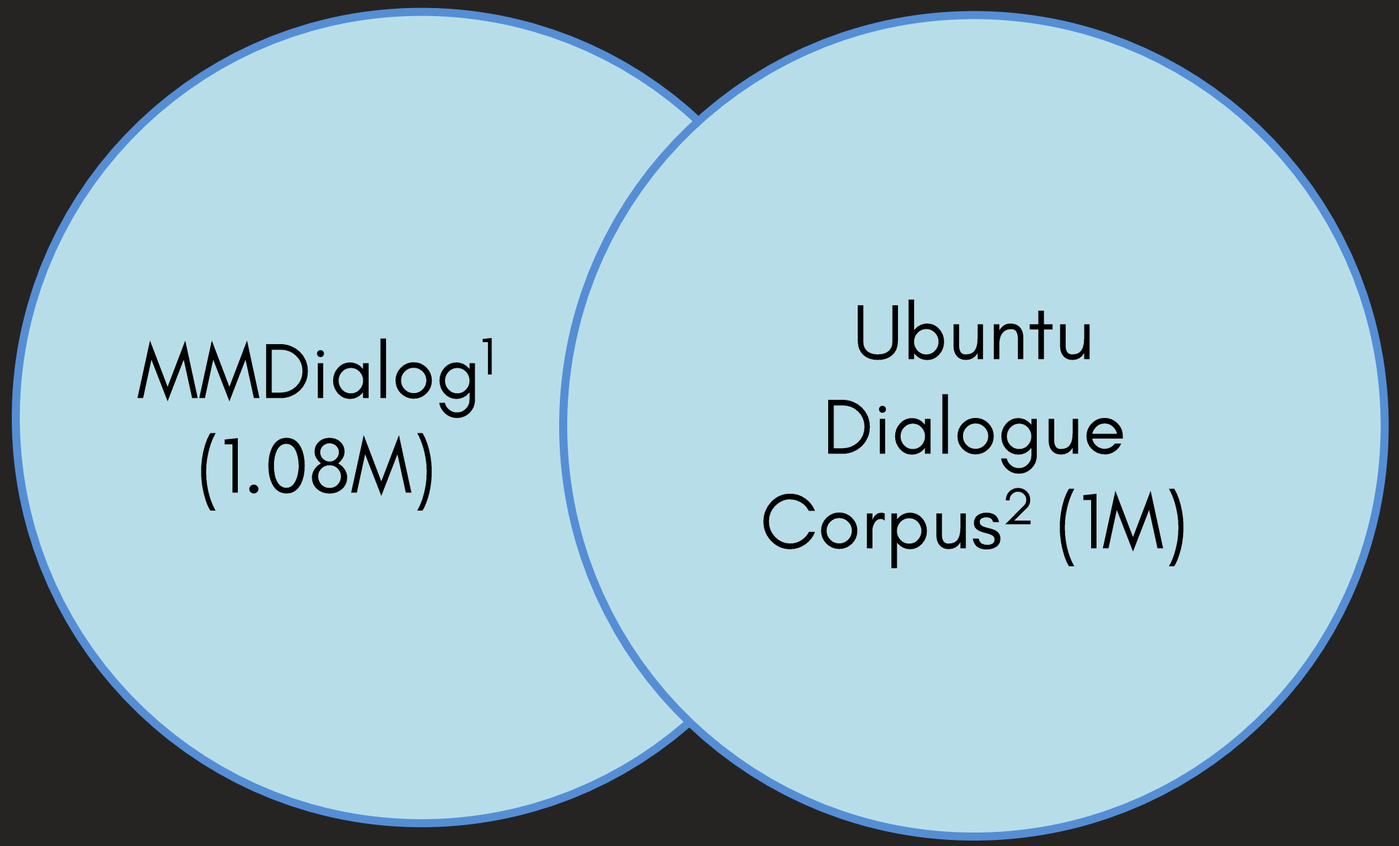
User learning from an AI generated character resembling an expert giving a lecture

Users becoming AI generated characters resembling historical characters for role playing

Pataranutaporn et al. 2021. AI-generated characters for supporting personalized learning and well-being. Nat Mach Intell.

Persona-Based LLM Agents: Issues

Dialogues Without Personas



Dialogues With Personas

MPCHAT[4] (15K)

ConvAI2[3] (17K)

1. Feng et al. 2023. MMDialog: A Large-scale Multi-turn Dialogue Dataset Towards Multi-modal Open-domain Conversation. ACL.
2. Lowe et al. 2015. The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems. SIGDIAL
3. Dinan et al. 2019. The Second Conversational Intelligence Challenge (ConvAI2).
4. Ahn et al. 2023. MPCHAT: Towards Multimodal Persona-Grounded Conversation. ACL.

Limited availability & diversity of data necessitate multiple training objectives

Limited availability & diversity of data necessitate multiple training objectives

- Supervised Finetuning + Contrastive Learning[1]
- Supervised Finetuning + Reinforcement Learning[2]

Learn Desirable Dialogues

Align With Preferences

1. Li et al. 2023. Learning to Know Myself : A Coarse-to-Fine Persona-Aware Training Framework for Personalized Dialogue Generation. AAAI.
2. Liu et al. 2020. You Impress Me: Dialogue Generation via Mutual Persona Perception. ACL.

Limited availability & diversity of data necessitate multiple training objectives

- Supervised Finetuning + Contrastive Learning[1]
- Supervised Finetuning + Reinforcement Learning[2]

Learn Desirable Dialogues

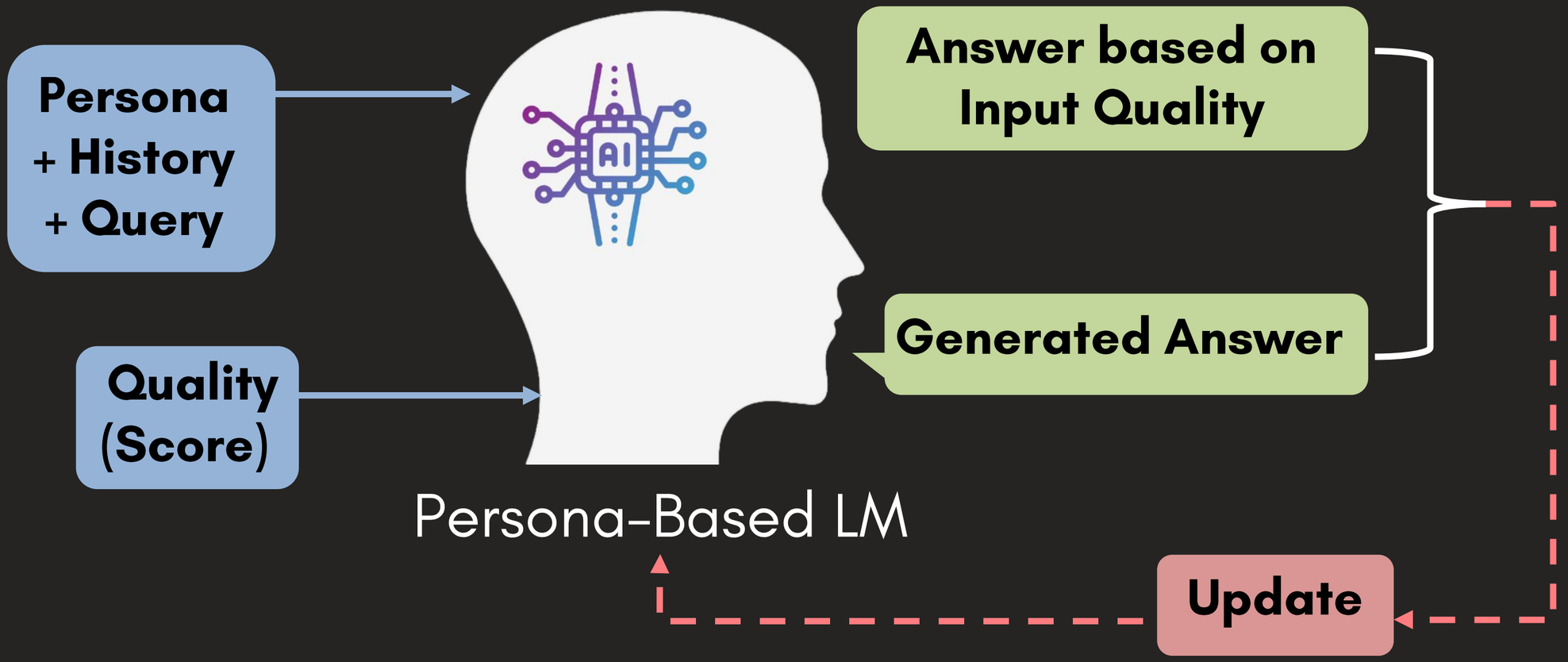
Align With Preferences

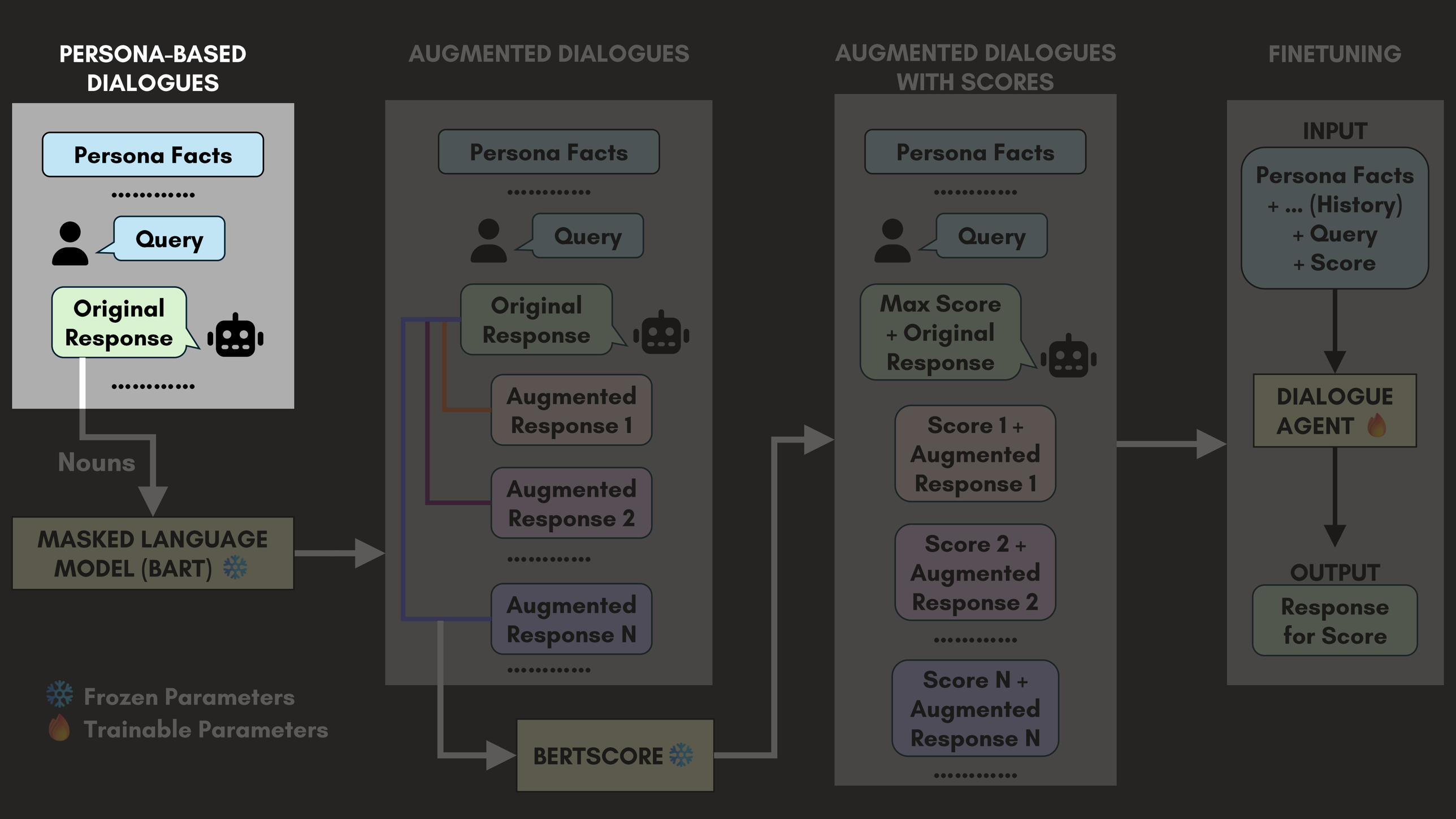
Not Ideal for Resource-Constrained Settings

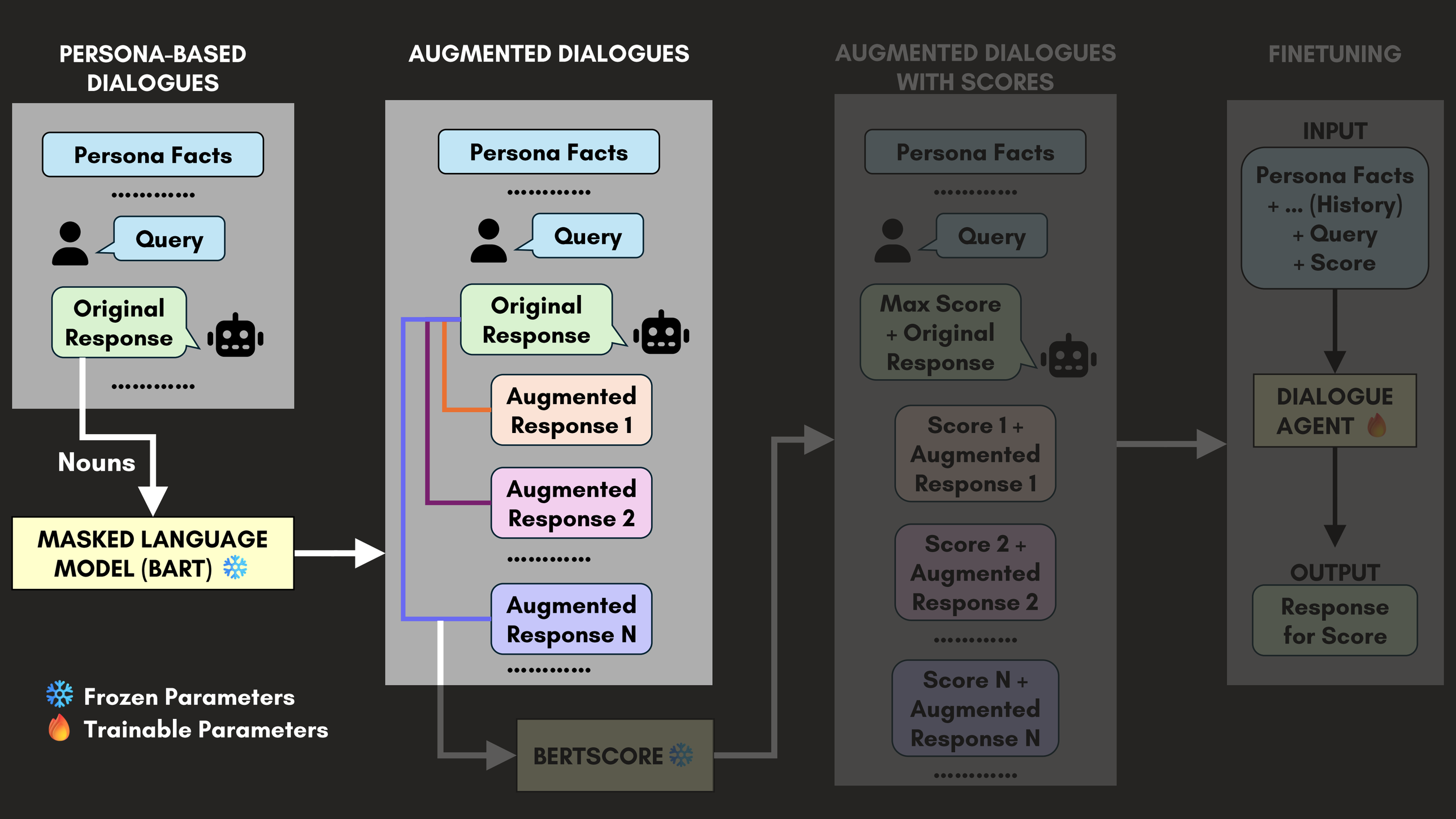
1. Li et al. 2023. Learning to Know Myself : A Coarse-to-Fine Persona-Aware Training Framework for Personalized Dialogue Generation. AAAI.
2. Liu et al. 2020. You Impress Me: Dialogue Generation via Mutual Persona Perception. ACL.

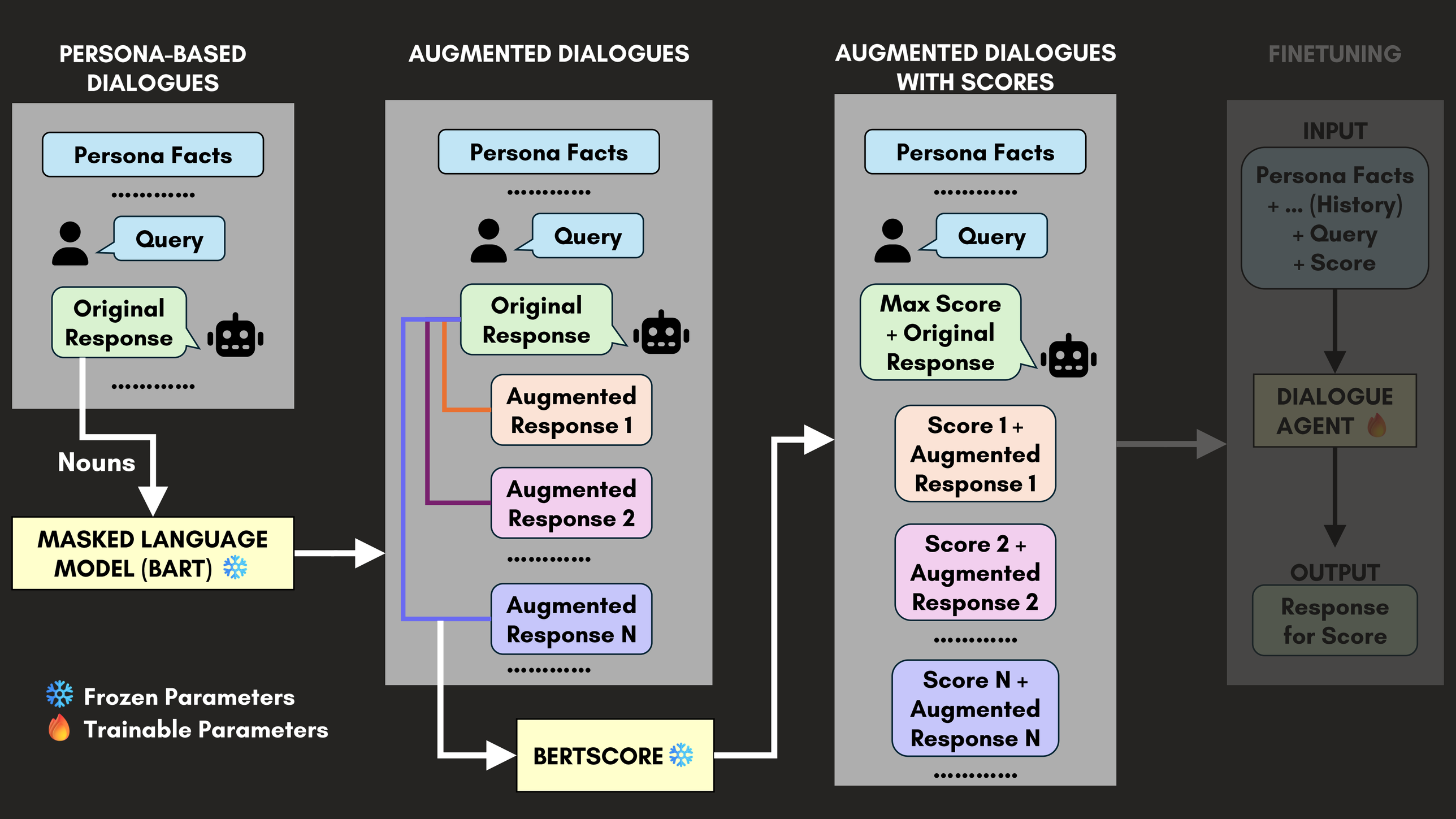
Can we unify finetuning & quality alignment into a single step?

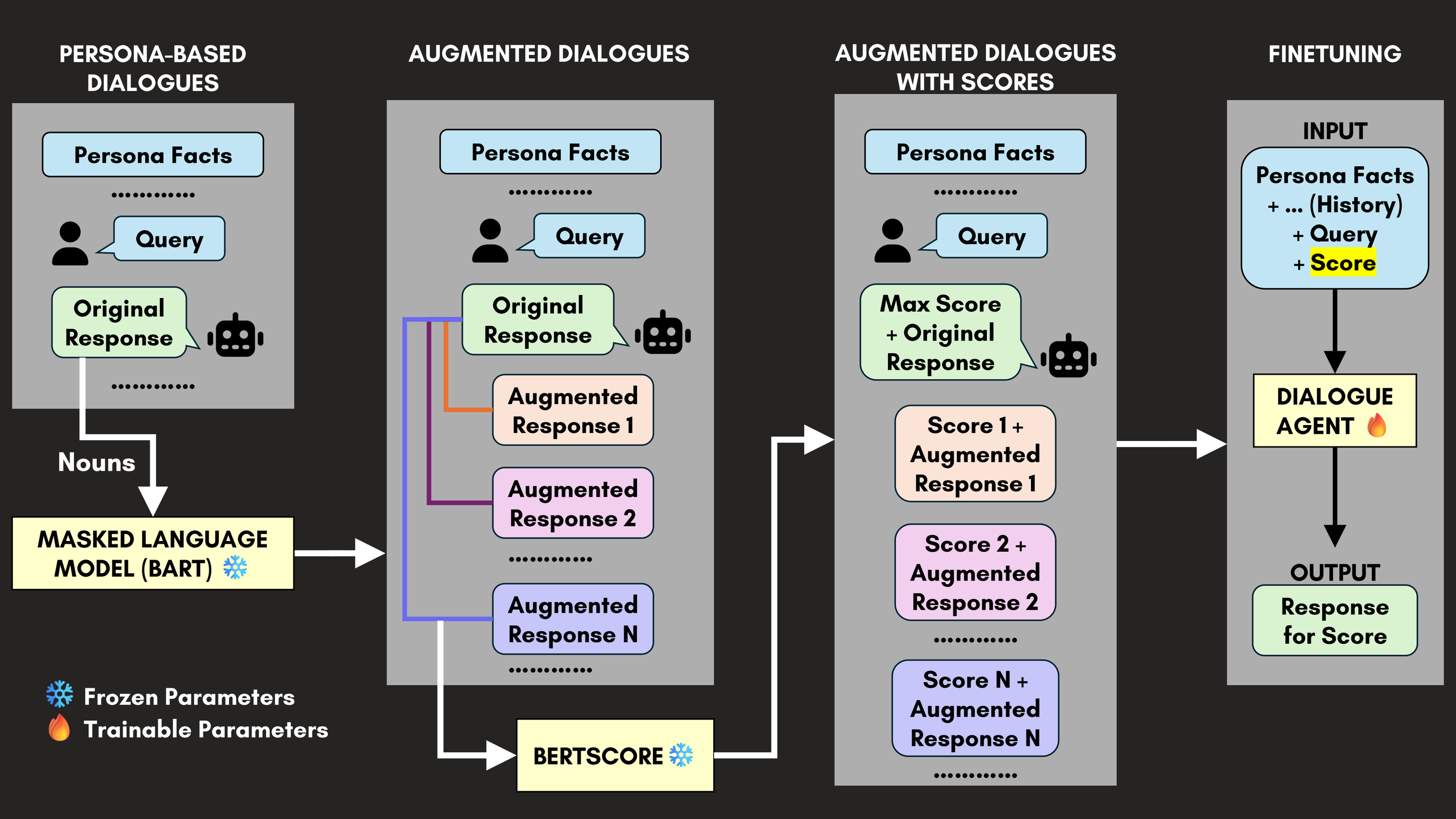
Score-Before-Speaking (SBS)




PERSONA-BASED DIALOGUES
Persona Facts
Query
Original Response
Nouns
MASKED LANGUAGE MODEL (BART)
AUGMENTED DIALOGUES
Persona Facts
Query
Original Response
Augmented Response 1
Augmented Response 2
Augmented Response N
BERTSCORE
AUGMENTED DIALOGUES WITH SCORES
Persona Facts
Query
Max Score + Original Response
Score 1 + Augmented Response 1
Score 2 + Augmented Response 2
Score N + Augmented Response N
FINETUNING
INPUT
Persona Facts + ... (History) + Query + Score
DIALOGUE AGENT
OUTPUT
Response for Score
Frozen Parameters
Trainable Parameters

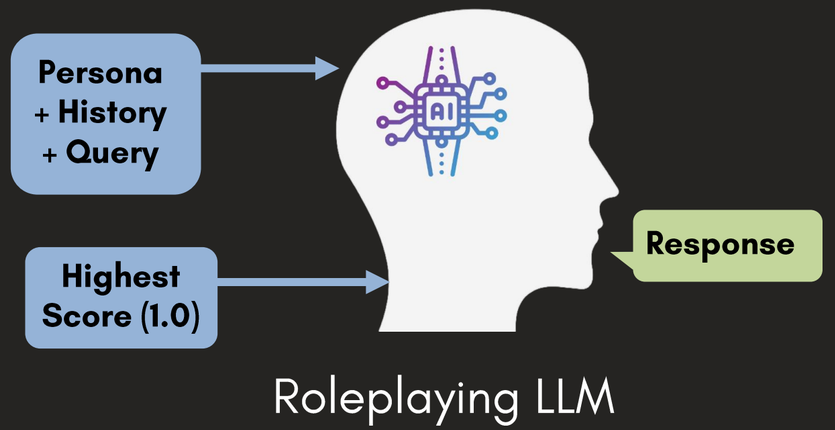

PERSONA-BASED DIALOGUES
Persona Facts
Query
Original Response
Nouns
MASKED LANGUAGE MODEL (BART)
AUGMENTED DIALOGUES
Persona Facts
Query
Original Response
Augmented Response 1
Augmented Response 2
Augmented Response N
BERTSCORE
AUGMENTED DIALOGUES WITH SCORES
Persona Facts
Query
Max Score + Original Response
Score 1 + Augmented Response 1
Score 2 + Augmented Response 2
Score N + Augmented Response N
FINETUNING
INPUT
Persona Facts + ... (History) + Query + Score
DIALOGUE AGENT
OUTPUT
Response for Score
Frozen Parameters
Trainable Parameters


PERSONA-BASED DIALOGUES
Persona Facts
Query
Original Response
Nouns
MASKED LANGUAGE MODEL (BART)
AUGMENTED DIALOGUES
Persona Facts
Query
Original Response
Augmented Response 1
Augmented Response 2
Augmented Response N
BERTSCORE
AUGMENTED DIALOGUES WITH SCORES
Persona Facts
Query
Max Score + Original Response
Score 1 + Augmented Response 1
Score 2 + Augmented Response 2
Score N + Augmented Response N
FINETUNING
INPUT
Persona Facts + ... (History) + Query + Score
DIALOGUE AGENT
OUTPUT
Response for Score
Frozen Parameters
Trainable Parameters


PERSONA-BASED DIALOGUES
Persona Facts
.....
Query
Original Response
.....
Nouns
MASKED LANGUAGE MODEL (BART)
AUGMENTED DIALOGUES
Persona Facts
.....
Query
Original Response
Augmented Response 1
Augmented Response 2
.....
Augmented Response N
.....
BERTSCORE
AUGMENTED DIALOGUES WITH SCORES
Persona Facts
.....
Query
Max Score + Original Response
Score 1 + Augmented Response 1
Score 2 + Augmented Response 2
.....
Score N + Augmented Response N
.....
FINETUNING
INPUT
Persona Facts + ... (History) + Query + Score
DIALOGUE AGENT
OUTPUT
Response for Score
Frozen Parameters
Trainable Parameters

Experiments

- Datasets: PERSONA-CHAT[1] & ConvAI2[2]
- Models: Llama 3.1[3] & DialoGPT[4]
- Inference using highest score



1. Zhang et al. 2018. Personalizing Dialogue Agents: I have a dog, do you have pets too? ACL.
2. Dinan et al. 2019. The Second Conversational Intelligence Challenge (ConvAI2).
3. Llama Team, AI @ Meta. 2024. The llama 3 herd of models.
4. Zhang et al. 2020. DIALOGPT : Large-Scale Generative Pre-training for Conversational Response Generation. ACL.

Evaluation

- Faithfulness to ground truth: Perplexity[1] and BLEU[2]
- Diversity of responses: Entropy-N[3] and Distinct-N[4]
- Persona-Consistency: C-Score[5]

1. Zhang et al. 2018. Personalizing Dialogue Agents: I have a dog, do you have pets too? ACL.
2. Papineni et al 2002. BLEU: a method for automatic evaluation of machine translation. ACL.
3. Zhang et al. 2018. Generating Informative and Diverse Conversational Responses via Adversarial Information Maximization. NeurIPS.
4. Galley et al. 2016. A Diversity-Promoting Objective Function for Neural Conversation Models. NAACL.
5. Madotta et al. 2019. Personalizing Dialogue Agents via Meta-Learning. ACL.

SBS outperforms/matches state-of-the-art

Subset of results. For the full evaluation, please refer to the paper.

	Method	Model	Size	BLEU-1 ↑	BLEU-2 ↑	BLEU-3 ↑	BLEU-4 ↑	Dist-2 ↑	Ent-2 ↑	C ↑
Small LMs	Baseline	DialoGPT	117M	17.99	8.62	4.91	2.84	15.67	6.56	60.84
	LTKM[1]	DialoGPT	117M	19.70	9.85	5.70	3.52	11.14	7.10	63.90
	PAA[2]	GPT2	254M	21.20	10.10	5.52	3.20	9.25	6.70	68.16
	SPT[3]	OPT	125M	20.63	9.86	5.56	3.22	17.38	7.29	65.13
	SBS (Ours)	DialoGPT	117M	22.11	11.30	6.75	4.21	16.52	7.10	70.60

Best **Second Best**

1. Li et al. 2023. Learning to Know Myself : A Coarse-to-Fine Persona-Aware Training Framework for Personalized Dialogue Generation. AAAI.
2. Huang et al. 2023. Personalized Dialogue Generation with Persona-Adaptive Attention. AAAI.
3. Huang et al. 2024. Selective Prompting Tuning for Personalized Conversations with LLMs. Findings of ACL.

SBS outperforms/matches state-of-the-art

Subset of results. For the full evaluation, please refer to the paper.

	Method	Model	Size	BLEU-1 ↑	BLEU-2 ↑	BLEU-3 ↑	BLEU-4 ↑	Dist-2 ↑	Ent-2 ↑	C ↑
Large LMs	Baseline	Llama 3.1	8B	21.57	10.94	6.43	3.88	22.43	7.87	64.07
	SPT[1]	Llama 2	7B	17.67	9.02	5.27	3.21	22.08	8.30	69.77
	LAPDOG[2]	T5	11B	22.48	11.41	6.73	4.14	24.76	8.15	59.92
	SBS (Ours)	Llama 3.1	8B	25.22	13.12	7.82	4.99	23.09	7.96	73.11

Best **Second Best**

1. Huang et al. 2024. Selective Prompting Tuning for Personalized Conversations with LLMs. Findings of ACL.
2. Huang et al. 2023. Learning Retrieval Augmentation for Personalized Dialogue Generation. EMNLP.

Scores correlate with Response Quality

Subset of results. For the full evaluation, please refer to the paper.

Score	BLEU-1 ↑	BLEU-2 ↑	BLEU-3 ↑	BLEU-4 ↑	Dist-2 ↑	Ent-2 ↑	C ↑
1.00	25.22	13.12	7.82	4.99	23.09	7.96	73.11
0.95	24.87	13.33	7.72	4.32	22.09	7.99	59.40
0.90	24.66	13.12	7.70	4.22	21.76	7.98	59.49
0.85	24.30	12.98	7.62	4.19	21.38	7.90	59.09
0.80	24.24	12.81	7.50	3.97	20.90	7.86	57.93
0.75	23.96	12.73	7.51	3.99	20.47	7.79	59.88

Expected Trend is Followed

Expected Trend isn't Followed

No change in Metric Value

Higher Score = Higher Persona-Consistency

PERSONA FACTS:
I'm a recording engineer and own my own studio.
My ex cheated and left me for a lawyer.
I am a 46-year-old single dad.
I live in California, but the recording artist market was dead last year.
I prefer being inside.

.....
USER: I wear a yellow construction hat, and **my dad is a lawyer**. Bummer, huh?

Score: 0.95

Score: 0.75

I'm not crazy about lawyers. What do you do for fun?

Yeah, I know a lot of people who do that.

CONTACT:
scasag@leeds.ac.uk

UNIVERSITY OF LEEDS

UK Research and Innovation

PROJECT PAGE →

Takeaways

- Score-conditioned training allows LMs to effectively delineate between persona-consistent and persona-inconsistent responses.

- Future work may include application to other natural language generation tasks or using scores for likelihood modification.

This research is part of a PhD project supported by United Kingdom Research and Innovation (Grant Reference: EP/S024336/1)