

Persona-based target-guided dialogue systems: A use case in dementia care

by

Renee Leung

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2023

© Renee Leung 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In recent years, conversational agents have shown potential in various applications. However, the development of conversational agents tailored for older adults, particularly those with age-related cognitive limitations, remains unexplored. Inspired by person-centred care, this thesis proposes a framework for building a persona-based target-guided conversational agent. First, we train a BART model to computationally extract elements of a user’s ‘persona’, which allows the agent to learn the individual’s background, preferences, and life stories. Then, we train a GPT-2 response generation model which leverages the extracted personas to generate personalised responses that preserve the identity of the individual, and train a keyword prediction kernel model to guide a conversation towards a given target topic. We evaluate the persona extraction model and response generation model on the public datasets PGDataset and ConvAI2 respectively. Then, we simulate conversations between a user agent and target-guiding agent to study the effects of incorporating persona information and keyword prediction into the conversational agent. Finally, we experiment the models on a new dataset constructed from a life-story interview transcript of older adults, and propose an experimental plan to evaluate the framework on real-life older adults. Our long term goal is to address the challenges faced by individuals with dementia and their caregivers by building proactive person-centred conversational agents.

Acknowledgements

I would like to thank my supervisor, Dr Jesse Hoey, for his continuous support, guidance, and encouragement throughout the process of completing my Master's. I am grateful to have been given the freedom and opportunity to explore different projects. Jesse's expertise in computer science, artificial intelligence, and social sciences has inspired many fruitful ideas in the interdisciplinary work.

I would also like to thank my readers, Dr. Pascal Poupart and Dr. Jim Wallace, for their time and valuable feedback on this thesis.

I would also like to thank my NEATO research team, Dr. Jesse Hoey, Dr. Heather Love, Heather Eustace, Jess Gano, and Alyssa Clarkson, for their useful discussions and guidance in shaping the future directions of the project. Thank you also to my fellow CHIL group members for their ideas and support.

Finally, thank you to my family and friends for their unconditional love and support throughout my academic journey.

Table of Contents

Author’s Declaration	ii
Abstract	iii
Acknowledgements	iv
List of Figures	vii
List of Tables	viii
1 Introduction	1
1.1 Thesis Outline	3
2 Background and Related Work	4
2.1 Dementia and Caregiving	4
2.2 Technology for Dementia Care	5
2.3 Conversational AI	6
2.4 Personalised Response Generation	8
2.5 Inferring Personal Information from Dialogues	9
2.6 Target-Guided Conversational Agents	11

3	Methodology	12
3.1	Persona Extraction from Dialogues	12
3.2	Persona-based Target-guided Response Generation	15
3.2.1	Generative Model	15
3.2.2	Kernel Generative Model	17
3.2.3	Persona Generative Model	18
3.2.4	Persona-Kernel Generative Model	19
3.2.5	Summary of models	20
4	Experiments and Results	21
4.1	Dataset	21
4.2	PGTask	23
4.3	Persona-based Target-guided Dialogue Generation	24
4.3.1	Simulation	25
4.3.2	Case Study	27
5	Application in Dementia Care	33
5.1	Experimentation with Older Adults Dataset	33
5.1.1	Persona Extraction	35
5.1.2	Persona-based Target-guided Response Generation	35
5.2	Moving Towards Dementia: Experimentation Plan for Older Adults	41
5.3	Limitations	42
6	Conclusion	44
6.1	Future Work	45
	References	46

List of Figures

2.1	Seq2seq model based on [64] with input sequence "BAD" and output sequence "CATS"	8
3.1	Proposed pipeline for building a persona-based target-guided dialogue system. First, the user's persona is extracted from an interview dialogue. Then, the agent interacts with the user based on the extracted persona, with the goal of guiding the conversation towards a target, e.g. <i>sleep</i> . Words that describe the persona are highlighted in green and transition keywords are italicised in purple.	13
3.2	Corrupted input transformations in BART based on [36]	14
3.3	Architecture of the GPT decoder based on [53]	16

List of Tables

3.1	Summary of models used in our experiments. The labels within the brackets indicate the corresponding section numbers where detailed explanations of each model can be found. We will compare the first four models as target-guiding agents in our simulation experiment. The self-persona generator is used as a simulated user agent.	20
4.1	Example of ConvAI2 dataset	22
4.2	Examples of PGDataset	23
4.3	Statistics of PGDataset	23
4.4	PGTask results on PGDataset	24
4.5	Evaluation of keyword prediction on the ConvAI2 test set	25
4.6	Evaluation of responses generated by each model. The score is computed using the cosine similarity between each response and context or persona. We compare the first four models for the role of the target-guiding agent. The self-persona generator is used as a simulated user agent.	26
4.7	Simulation results. We compare different target-guiding models where the self-persona generator is used as the user agent.	28
4.8	Good example of simulated conversations guided by the kernel model and persona-kernel model. Predicted next keywords by the target agent are italicised within square brackets.	30
4.9	Average example of simulated conversations guided by the kernel model and persona-kernel model. Predicted next keywords by the target agent are italicised within square brackets.	31

4.10	Failed example of simulated conversations guided by the kernel model and persona-kernel model. Predicted next keywords by the target agent are italicised within square brackets.	32
5.1	Interview questions of our older adults dataset	34
5.2	Statistics of our older adults dataset	35
5.3	Examples of extracted personas in an interview transcript with Mrs. R . . .	36
5.4	Examples of extracted personas in an interview transcript with Mr. M . . .	37
5.5	Examples of simulated conversations between Mrs. R and the target agent	39
5.6	Examples of simulated conversations between Mr. M and the target agent	40
5.7	Erroneous example of persona generation	42

Chapter 1

Introduction

Dialogue systems, or conversational agents, have gained significant attention in recent years due to advancements in natural language processing and deep learning. To bridge the gap between dialogue systems and human conversational capabilities, there is a need to incorporate personal information and behavioural cues into conversational agents, so that agents can better adapt to users' current states and make the interaction more engaging.

More recently, persona-based dialogue systems were introduced to make the conversational experience more personalised for users. Zhang et al. [79] introduced the **PersonaChat** dataset, in which each speaker is delegated a set of profile sentences that describe their individual attributes and preferences (*persona*). The authors conducted an experiment involving human Turkers engaging in conversations with agents trained both with and without personas. The Turkers evaluated the responses from each model on 100 dialogues and their results demonstrated that conversational agents that incorporate user persona information are significantly more engaging and consistent compared to those that do not [79].

Personalised dialogue systems can be utilised by a range of applications, including recommendation systems [62, 82, 71, 38], automated customer service [2], healthcare [19, 20, 43, 44, 9] and more. In this thesis, we will explore a potential use case in dementia care. Dementia is a neurological condition which impacts the cognitive function of older adults, leading to difficulties in thinking and memory. Individuals with dementia may experience a gradual fragmented loss of their personal history, social roles, and an overall sense of self-identity. Studies have shown that increased social interactions are beneficial in decreasing cognitive decline in older adults [3]. However, communication with people with dementia can be difficult due to variations in their behaviour. Caregivers often adjust

their communication style or prepare cues to stimulate their memory and reinforce aspects their identity [18]. Therefore, it is beneficial to acquire information about their past memories, life stories, and preferences to better connect with them, which helps improve the relationship between the person with dementia and the caregiver.

Furthermore, individuals with dementia may exhibit impaired judgement and harmful behaviours which affect their health and safety [61]. Family members often have to take on the role of informal caregivers to support the individual’s wellbeing. Dementia caregiving, which encompasses providing support for the individual’s everyday activities like bathing or eating, can pose emotional, mental, and physical burdens on the caregiver when the individual refuses to participate in these necessary tasks. To address these challenges, caregivers or counsellors often initially employ strategies that involve aligning with the person’s perceptual and emotional space to establish a connection for effective communication. Once trust is established, the counsellor can gradually introduce practical changes in thinking and behaviour. This approach, which is commonly applied in therapy for depression or addiction, recognises the whole person and places them in a supportive environment, thereby encouraging adaptive behaviours that the person is more likely to adopt.

Given the forthcoming surge of elderly individuals from the baby-boomer generation and the anticipated scarcity of young caregivers, there is a compelling motivation to harness computer systems to address the needs of people with dementia and their caregivers. To provide direction in building a labour-effective and personalised system for this purpose, we propose to build a persona-based target-guided conversational agent. It is *persona-based* such that it incorporates information about the user’s preferences and significant life moments, allowing the conversational agent to engage with the individual while preserving their personal identity. It is also *target-guided* such that the agent smoothly transitions the conversation from any topic to a target keyword, in hopes of aligning the user’s cognitive perspective to the context of a desired task or behaviour. It is important to emphasise that while this agent shows potential for application in dementia care, this work represents an initial step towards achieving that goal.

This thesis will study two main tasks: inferring persona information from dialogues and building a persona-based target-guided conversational agent. In the former task, we will experiment generating personas using the BART (Bidirectional and Auto-Regressive Transformers) model [36] on the Profile Generation Dataset (PGDataset) [57], which is shown to exceed the dataset’s baseline performance. In the latter task, we will train a response generative model which (1) smoothly transitions towards a target keyword while remaining contextually relevant to the dialogue history, and (2) is engaging to the user by maintaining alignment with the user persona. This is achieved by training a

keyword prediction kernel and a conditional keyword response generative model which leverages persona information on the ConvAI2 dataset [14]. We compare the effectiveness of incorporating both the keyword prediction kernel and persona information based on context relevance, persona similarity, and the success rate of reaching the target. We hypothesise that adding persona information to the agent would increase the success rate of reaching the target and decrease the average number of turns to reach the target. This is because the keywords selected will be closer to the user’s persona, hence yielding smoother transitions assuming the user often talks about their own persona. Our experiments demonstrate that encompassing both components, the user persona and keyword prediction kernel, can help reach the target at a higher success rate. Finally, we experiment both models on a podcast dataset featuring older adults, and present a future plan to evaluate the framework’s applicability in real-life interactions involving older adults. This roadmap aims to advance our methodology towards addressing the needs and characteristics of individuals with dementia.

1.1 Thesis Outline

The thesis is organised as follows:

- Chapter 2 discusses the background and related works for dementia care and conversational systems.
- Chapter 3 describes the methodology of the proposed framework, including inferring persona information from dialogues and generating persona-based target-guided responses.
- Chapter 4 presents experimental results of the methodology on the PGDataset and ConvAI2 datasets.
- Chapter 5 applies the methodology on an older adults dataset and proposes an experimental plan involving real-life interactions with older adults.
- Chapter 6 includes the conclusion and future work for this study.

Chapter 2

Background and Related Work

2.1 Dementia and Caregiving

Currently, more than 55 million people suffer from dementia in the world and the figure increases by nearly 10 million every year [47]. Dementia is a permanent neurological disorder in older adults caused by damage to neurons. It is characterised by cognitive decline which gradually worsens over time. The most common types, including Alzheimer's, cause difficulty in thinking, language, memory loss, and other human activities that affect the individual's daily functioning. As dementia progresses, the individual may experience a loss of sense of self. This sense of self is defined by their overall experience of personhood, including their memories, history, and individuality. As a result, it may be challenging for older adults with dementia to maintain a consistent identity, accurately interpret social interactions and norms, or have a dialectical relationship with society [46].

Identity may be viewed as a multidimensional concept characterising an individual's traits, tastes, habits, socio-demographic backgrounds, and social roles [68]. Identity may also be constructed through interacting with others and the environment. This notion of identity is in line with Kitwood's person-centered theory and malignant social psychology [34], which claims that the sense of self is heavily affected by the attitudes of others, rather than the disease itself [59, 58]. Effectively, caregivers are encouraged to look from the perspective of the person with dementia and communicate with identity cues to stimulate social interactions, helping them preserve self-perception and maintaining their identity [63, 68]. For instance, reminiscence therapy can stimulate memory about people, places, and significant events in their lives [6]. A study finds that reminiscence about positive

experiences would restore the saliency of the demented person’s former identity [16]. Thus, knowing the person is critical to a trusting and fruitful relationship in dementia care.

In addition to identity challenges, older adults with dementia may exhibit various maladaptive behaviours, including anger, poor hygiene, and poor judgement [61]. Anger can be expressed through outbursts, aggression, or irritability, where poor hygiene practices may include neglecting personal grooming and cleanliness. These behavioural changes often pose substantial challenges for caregivers and impact the individual’s overall well-being. Caregivers are often confronted with a dilemma of prioritising the individual’s autonomy versus safeguarding their safety and wellbeing. For example, when an older adult with dementia refuses to follow caregiver instructions, such as resisting bathing, eating, or going to bed, whether the caregiver should respect their desire even if it entails potential harm to their health, or override their preferences to enforce necessary actions, both options present suboptimal outcomes. Skillful communication that connects with the person with dementia is crucial for providing appropriate care and reinforcing their dignity and quality of life.

Within the United States alone, over 11 million unpaid family caregivers devote 18 billion hours to care for people with dementia, where 95% of them reported high or very high levels of emotional stress associated with their caregiving responsibilities [1]. These burdens not only have detrimental effects on the relationship between the demented person and the caregiver, but also pose potential risks of anxiety, depression, and cardiovascular diseases [10]. With the increasing number of people developing dementia, it motivates us to leverage technology to alleviate the mental strain experienced by caregivers.

2.2 Technology for Dementia Care

In recent years, there has been significant research on the development of assistive tools for dementia care to address the challenges faced by individuals living with dementia and their caregivers. By harnessing technology, these solutions are designed to provide accessible and personalised support for people affected by dementia socially and cognitively.

Early works introduced robot companions to satisfy the need for interactions with people with dementia. PARO is an animal-therapy inspired social robot which uses reinforcement learning to adjust its behaviour according to user feedback, such as stroking (positive) or beating (negative), to improve the mood, physiological condition, and social activity levels of older adults and people with dementia [70, 69].

From the cognitive approach, NAO and JESSIE are social robots which can be programmed to simulate emotions and accommodate personalised specifications. They have been adapted in cognitive training interventions for older adults with cognitive impairment, including memory training sessions and neurorehabilitation programs, in which participants showed noticeable improvements in attention levels and a reduction in depressive symptoms [51, 35].

Furthermore, immersive technology has been developed to provide reminiscence therapy to stimulate sensory perceptions, facilitate memory recall, and preserve identity. Goodall et al. conducted a study in which they gathered user profiles of individuals with dementia and reconstructed personalised narrative experiences within **SENSE-GARDEN** through interactive photographs, landscapes, scents, and sounds that were tailored to the individual’s life story. Through qualitative analysis, the authors observed positive outcomes, including stimulation of positive emotions, preservation of identity, and improved interpersonal relationships among the participants [22].

Other studies have shown that despite encountering limitations, such as occasional misunderstandings, older adults and people with dementia indicate positive emotional responses and interest when engaging with social robots [52, 21]. Moreover, Eustace found that technology that incorporates personally relevant elements, such as autobiographical narratives, further enhances older adults’ interests in engaging with it [17]. In a similar vein, caregivers also share affirmative feedback on the effectiveness of technological support tools in meeting caregiving needs [5]. These findings highlight the potential benefits of incorporating social systems in dementia care and motivate us to build personalised assistive technology that is tailored to meet individual needs and create meaningful experiences.

Inspired by person-centred care, this thesis introduces a persona extraction model and a personalised target-guiding conversational agent as a foundational framework. The intended trajectory is to apply this framework to dementia care, so that future agents can interact with individuals with dementia by comprehending their preferences and maintaining their identity, while skillfully guiding them to encourage their interest and ability to complete daily tasks.

2.3 Conversational AI

A dialogue system is a computer-based system that interacts with humans through natural language. These systems aim to engage in conversations and automatically generate responses that mimic human communication patterns. Dialogue systems are generally separated into two types, task-oriented and chit-chat systems. On one hand, task-oriented

dialogue systems focus on carrying out specific tasks, such as booking a flight ticket or scheduling a call, and are designed to recognise user intents. On the other hand, chit-chat systems focus on chatting and engaging the user in a conversation by generating responses that are diverse and relevant to the dialogue context. In this work, we study open-ended chit-chat conversational agents.

Traditional conversational AI primarily relied on accessing structured data stored in databases and rule-based methods to model dialogue generation. ELIZA was the first chatbot introduced by Weizenbaum in 1966, which uses pattern-matching and transformation rules to produce responses [73]. In 1972, another chatbot PARRY was developed by Colby [12]. PARRY uses a similar rule-based approach like ELIZA but incorporates a mental model to simulate affective variables including anger, fear, and mistrust. However, these approaches often faced challenges in handling more complex, diverse, unstructured user inputs, including raw text, and had limitations in scalability.

In recent years, deep-learning approaches have attracted much attention due to their remarkable success in computer vision and natural language processing tasks. These models are trained from large amounts of unstructured data to learn important features for dialogue understanding and generation. Dialogue systems can be further categorised into retrieval-based and generation-based approaches. Retrieval-based systems rank a given pool of candidate responses in a dataset usually through computing similarity measures between features of the dialogue context and each candidate, then choose the top candidate as the response. While these models are good at providing factual responses, they may lack the ability to output appropriate responses in an unseen context. On the contrary, generation-based systems produce responses by outputting a sequence of tokens that may not be observed in the dataset. These systems are usually modelled by a sequence-to-sequence (seq2seq) encoder-decoder.

The original seq2seq model proposed by Sutskever et al. [64] employs a Long Short-Term Memory (LSTM) architecture [27], where an LSTM encodes an input sequence into a fixed-length vector representation, then another LSTM decodes it into an output sequence (Figure 2.1). More recently, transformers have been shown to surpass LSTMs in terms of speed and performance [67]. The transformer model utilises the self-attention mechanism to model contextual relationships between words within a sequence. This feature is especially valuable in dialogue systems, as it enables the capturing of long-range dependencies. Notably, transformer-inspired large language models (LLMs) have demonstrated remarkable capabilities in generating responses that exhibit broad knowledge and human-like conversational qualities.

In this work, we focus on generation-based models due to their better performance and

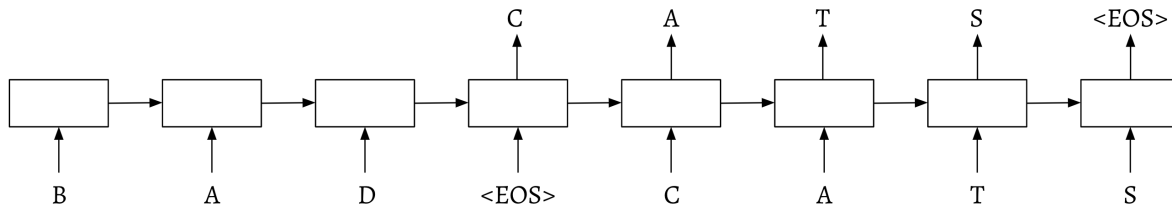


Figure 2.1: Seq2seq model based on [64] with input sequence "BAD" and output sequence "CATS"

flexibility to apply to the dementia care domain.

2.4 Personalised Response Generation

Incorporating personal information into dialogue systems is important for generating responses that emulate human qualities, such as ensuring consistency in personality and topic interests. Earlier works delved into creating an identity profile for the chatbot and generating conversational text based on its given profile attributes (*self-persona*). Li et al. [37] first proposed a persona-based conversation model which incorporates speaker embeddings into the response decoder. Similarly, Zheng et al. [83] integrate personal traits (age, gender, and location) as key-value pairs into the response generation process by encoding the persona information into a vector representation using a trait fusion model then decoding it in the decoder output layer. Zhang et al. [79] introduce a persona dialogue dataset where each speaker is labelled a set of profile sentences describing their occupation, hobbies, family, etc. Their baseline experiments find that models trained with conditioned self-personas generate responses which have higher engagingness and consistency compared to those trained without personas. Chen et al. [8] model persona and dialogue context independently by learning an external memory which formulates persona-response entailment relations as a premise-hypothesis task and a similar internal memory for dialogue discourse.

Later studies found that humans adapt their speaking patterns not only according to their own characteristics but also based on the attributes and identity of their conversational partner. Chen et al. [7] show that the partner’s gender, a typical identity variable, influenced the responder’s language by successfully classifying responses into each gender and analysing corresponding topic keywords on Chinese social media data. Gu et al. [24]

empirically demonstrate that utilising partner personas improved accuracy for response selection. Several models have been proposed to incorporate the partner’s identity profile into the conversational agent (*other-persona*). Dong et al. [15] study the user-aware response generation task on a microblog dataset. They use the Context-aware Knowledge Pre-selection (CaKe) model [81] to select the most relevant profile for the given dialogue context, then based on a probability, generates tokens from the vocabulary distribution or from copying tokens from the selected profile. The authors also propose a meta-learning method to learn sparse user personas as different tasks. Cho et al. [11] develop a user-personalised response generation model using conditional variational inference by introducing a latent variable which detects implicit partner persona. Lu et al. [42] train a reinforcement learning agent to jointly extract partner persona and generate response conditioned on the partner persona.

Personalised dialogue generation has been explored in a range of social influence applications which influence users’ states such as thoughts and behaviours. Mousavi et al. [44] train a response selection model based on user personal narratives to carry out follow-up dialogues in a Cognitive Behavioural Therapy setting. In their experiments, they show that their model achieves a level of quality that is on par with the responses delivered by experienced psychotherapists. Zhong et al. [84] and Cheng et al. [9] leverage user persona to generate personalised empathetic responses for an emotional support system. Their analysis reveals that an increase of user persona mentions lead to responses of higher empathy and a greater decrease in emotional intensity. Finally, Wang et al. [72] investigate the influence of user demographics and personality traits on persuasive strategies used in a dialogue system aimed at encouraging charity donation. These encouraging outcomes serve as a strong motivation to apply personalised dialogue generation methods in the domain of dementia care.

2.5 Inferring Personal Information from Dialogues

Many personalised dialogue generation models require predefined personas for the agent and/or the user. Recognising the persona of the conversational partner is important for understanding their intentions and responding with language that is better suited to enhance the receptiveness of the conveyed information. However, persona information may not be readily available or could be expensive to obtain in real-world applications. To reduce human effort, a model can be trained to computationally extract persona information from dialogues. Here, persona information includes individual biographical backgrounds, preferences, life stories, etc.

Earlier works include extracting biographical information from transcripts using entity recognition, relation detection, event detection, and coreference resolution techniques [30], or classifying gender, age (below or above 30), regional origin (Eastern or Southern India), and political orientation as binary attributes [55]. However, these techniques may struggle to handle more complex nuances and may not capture the full dimensions of the speaker’s persona.

Using graph approaches, Mousavi et al. [45] utilise entity-relation extraction techniques to build a personal narrative graph from patient life events in psychotherapy sessions. Yen et al. [78] directly extract life events as *subject-predicate-value* triples from Twitter texts to build a personal knowledge base. However, the list of possible values can be infinitely long, or values may not be explicit in the dialogue text. To address this problem, Wu et al. [75] first use a multi-label classifier to predict appropriate predicates associated with a given utterance, then identify the corresponding subjects and objects from the utterance using an entity generator. Similarly, Tiginova et al. [66] use an information retrieval approach to search for the most appropriate value from Wikipedia documents. However, these methods are limited to a set of relation types and are unable to capture the comprehensive aspects of a person’s persona.

More recent works attempt to predict personal information from conversations as persona statements, which contain richer information than key-value pairs. Gu et al. [23] design the Speaker Persona Detection task to select the persona which best matches the conversational context using a matching network. However, this approach relies on a pre-defined list of candidates, which may not reflect real-life personas accurately. To tackle the problem of extracting persona sentences from unseen data, Ribeiro et al. [57] reconstruct the **PersonaChat** dataset to introduce the Persona Generation Task (PGTask), which aims to generate a profile sentence conditioned on a given utterance.

There has also been research on constructing latent representations of user persona. Li et al. [37] encode speaker embeddings from Twitter data and television transcripts, in which they are able to cluster users based on their backgrounds such as age and location. Cho et al. [11] model user persona from dialogue context implicitly through conditional variational inference. Their results also show clusters of different topics in the latent space of the trained posterior network. Since these methods are difficult to interpret, we focus on extracting persona sentences in this work. In particular, we will study the existing PGTask.

2.6 Target-Guided Conversational Agents

As mentioned in the above section, many social influence applications strive to achieve specific objectives which influence user behaviour, ranging from facilitating therapeutic sessions in psychotherapy, promoting healthy behaviour change, to making recommendations. Tang et al. [65] introduce the concept of Target-Guided Open-Domain Conversation, in which the conversational agent takes a proactive role in directing the conversation towards a target topic that is unknown to the user. The goal is to make smooth transitions towards the target while naturally chatting with the human.

Wu et al. [76] approaches the problem by utilising knowledge graphs to construct a path towards the target knowledge entity. In a similar vein, Sevegnani et al. [60] and Gupta et al. [25] bridge the current topic to the new topic in a single turn transition utterance by employing reasoning techniques within commonsense knowledge graphs.

More recently, Kishinami et al. [33] introduce the task of Target-Guided Open-Domain Conversation Planning, where the agent plans the full sequence of self-conversing turns to reach the target. However, these methods do not consider the user persona and therefore may not keep users engaged in new topics. Wang et al.’s work [71] is most related to our work, wherein they build a target-guided recommendation system grounded on the user profile and domain knowledge. They achieve this by planning a path sequence of dialogue actions and topics. Their work is different from ours in that recommendation dialogue actions are task-oriented rather than chit-chat based.

Chapter 3

Methodology

The proposed pipeline begins with extracting the user’s persona from a dialogue, followed by generating target-guiding responses in a conversation conditioned on the inferred user persona (Figure 3.1).

3.1 Persona Extraction from Dialogues

We train our persona extraction model using the setup outlined in PGTask [57]. The notion of ‘persona’ used throughout the thesis refers to a characteristic profile associated with an individual, which contains information about their background, preferences, hobbies, values, etc. The goal of the persona extraction task is to infer information pertaining to the persona of an individual from their utterances. Formally, given an utterance $u = \{w_1^u, \dots, w_m^u\}$, the goal is to generate profile sentence $p = \{w_1^p, \dots, w_k^p\}$ that is associated with u , where w_i represents a sequence of words. In the original paper, the authors develop a generative model as a baseline that maximises the log-likelihood over p :

$$\mathcal{L} = \sum_{i=1}^k \log P(w_i^p | w_1^u, \dots, w_m^u, w_1^p, \dots, w_{i-1}^p).$$

However, our experiments find that simply using transformer decoders, including GPT-2 [54] as used in the paper, struggle to handle noisy data. In conversational text, noise can include incomplete sentences, repetitive phrases, improper grammar, and ambiguity. These types of noise are not typically observed in the training data, making it challenging to apply the trained model to other domains where such noise is more prevalent. Considering

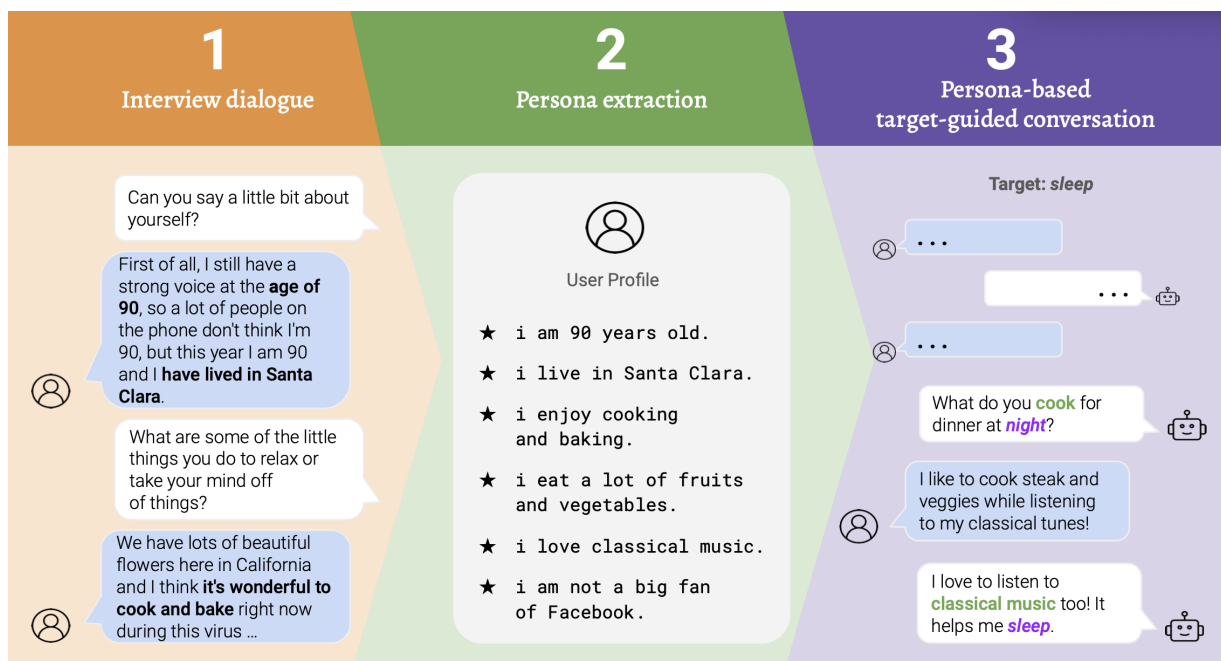


Figure 3.1: Proposed pipeline for building a persona-based target-guided dialogue system. First, the user’s persona is extracted from an interview dialogue. Then, the agent interacts with the user based on the extracted persona, with the goal of guiding the conversation towards a target, e.g. *sleep*. Words that describe the persona are highlighted in green and transition keywords are italicised in purple.

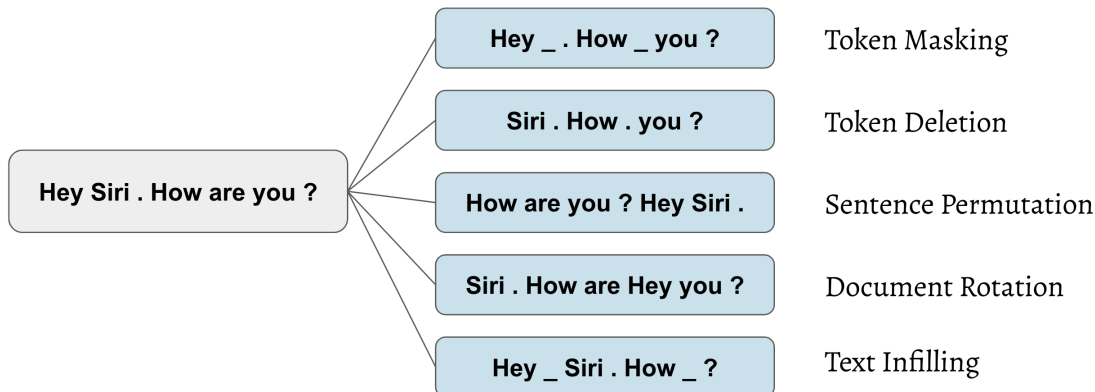


Figure 3.2: Corrupted input transformations in BART based on [36]

the linguistic characteristics commonly found in the language of older adults, which often involves fragmented sentences and the use of filler words such as "ums" [48, 29], we choose to train an alternative model that can better handle noisy inputs.

Specifically, we propose to train BART [36] as our generative model. BART is a denoising pretrained seq-to-seq encoder-decoder model which has been successfully employed in a range of NLP tasks, including text classification, generation, translation, and comprehension. BART is pretrained on a large text corpus with corrupted inputs to reconstruct the original text. Its architecture combines the strengths of BERT [13] (bidirectional encoder) and GPT [53] (autoregressive decoder). The bidirectional encoder learns to understand and encode corrupted data by processing surrounding words (both the left and right directions), while the autoregressive (left-to-right) decoder learns to decode and generate masked embeddings based on the encoded information from the encoder. The incorporation of diverse corrupted transformations, including token masking, token deletion, sentence permutation, document rotation, and text infilling (Figure 3.2), allows the model to be robust to noise, hence making it a suitable candidate for the persona extraction task when working with text data from older adults.

Consider $C(x)$ to be the corruption function, $e_\phi(\cdot)$ the encoder, and $d_\theta(\cdot)$ the decoder. In the pretraining step, BART learns to denoise corrupted texts and recover them in an unsupervised manner by minimising the reconstruction loss, that is, the cross-entropy

between the decoder output and the original input:

$$\ell(x, y) = - \sum_t \sum_{j=1}^n x_j^t \log P(y_j^t = 1|x),$$

where t is a time step which runs over the length of the text, $P(y|x) = \text{softmax}(y)$, $y = d_\theta(e_\phi(C(x)))$ is the unnormalised logits of the decoder output, x is the one-hot vector of the original document, n is the vocabulary size, and ϕ, θ are trainable parameters.

To fine-tune the pretrained BART model on our supervised persona extraction task, we minimise the cross-entropy loss between the model output and target profile statement:

$$\ell(p, y) = - \sum_{i=1}^k \sum_{j=1}^n \mathbf{w}_{ij}^p \log P(y_j = 1|w_1^u, \dots, w_m^u, w_{<i}^p),$$

where \mathbf{w}_i^p is the one-hot vector of w_i^p .

3.2 Persona-based Target-guided Response Generation

We adopt the target-guided conversation framework by Tang et al. [65] as the foundation for our model. Instead of using a response retrieval strategy, we will train a generative model. Also, we will explore whether there are benefits of incorporating persona information into the model.

Consider a conversation history defined by a sequence of utterances $\{u_1, \dots, u_n\}$ by the user and the agent, and a set of persona statements $\{p_1, \dots, p_m\}$ of the user. The agent aims to generate a target response r that (1) smoothly transitions towards a target keyword while remaining contextually relevant to the dialogue history, and (2) is engaging to the user by maintaining alignment with the user persona.

Below we describe the models that form the foundation of our proposed model and serve as baselines in our experiments.

3.2.1 Generative Model

In the generative model, the persona information or the target are not taken into consideration. The model solely generates a response that is contextually relevant based on the

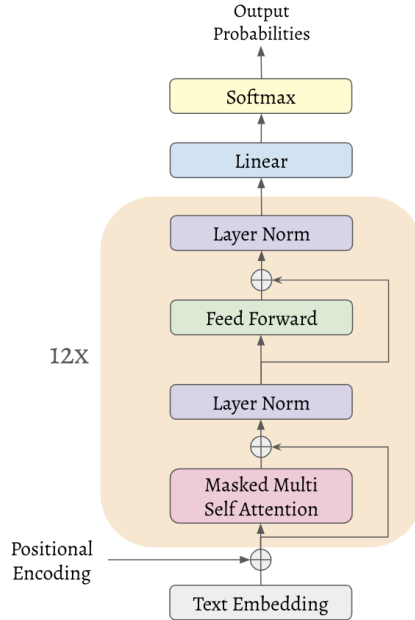


Figure 3.3: Architecture of the GPT decoder based on [53]

conversation history. The training objective is formalised as:

$$\mathcal{L}_{\text{gen}} = -\frac{1}{N} \sum_{t=1}^N \log P(r_t | u_1, \dots, u_n, r_{<t}),$$

where N is the length of the response r . In order to compute $P(r|u_1, \dots, u_n, r_{<t})$, we propose fine-tuning the transformer decoder GPT-2. GPT-2 is a language model which is pretrained with a large number of parameters to write coherent, human-like text. Taking an embedding as input, it generates tokens in an autoregressive fashion. The GPT-2 model has the same architecture as its predecessor GPT, which is almost identical to the decoder part of the Transformer model introduced by Vaswani et al. [67]. The GPT architecture is shown in Figure 3.3, where the model outputs probabilities of the next token to be generated over all words in the vocabulary.

In our model, the input representation is constructed by the sum of the token embeddings, positional embeddings, and role embeddings. The token embeddings are computed using trainable weights of the word token embeddings of the GPT-2 model. The positional embedding ranges from zero to the sequence length to denote the order of tokens, and

the role embedding distinguishes the utterances by the user and agent. Additionally, each utterance is encapsulated by a [BOS] (beginning of sequence) and [EOS] (end of sequence) token. During inference, the decoder stops generating once the [EOS] token is predicted.

3.2.2 Kernel Generative Model

The kernel generative model aims to smoothly transition towards the target in a conversation. Following Tang et al. [65], we use a turn-level keyword transition approach to generate a contextually relevant response that steps closer to the target. Specifically, a keyword prediction kernel is trained to score all candidate keywords for the next utterance.

Keyword extraction During both the training and inference stages, all nouns, verbs, and adjectives are extracted and lemmatised (e.g. "eating" becomes "eat") using the NLTK toolkit [4]. Each word is then assigned a score using TF-IDF, and if the score reaches a threshold, the word is extracted as a keyword.

Kernel scoring Originally proposed by Xiong et al. [77], given the keywords of the current turn $\{w_1, \dots, w_h\}$ and a candidate keyword c for the next turn, the kernel ranking model first computes a similarity vector v between the normalised word embeddings and the candidate keyword embedding:

$$v_i = \text{cos-sim}(z_{w_i}, z_c),$$

where z_w is the trainable word embedding of w . Then, the similarity scores are fed into K RBF kernels to be pooled into K -dimensional feature vector:

$$\mathcal{K}_k(v) = \sum_{i=1}^h \exp\left(-\frac{(v_i - \mu_k)^2}{2\sigma_k^2}\right),$$

where μ_k and σ_k^2 are the mean and variance of the k -th kernel. Essentially, \mathcal{K}_k determines the distribution of similarity scores around μ_k , hence the more similar the candidate keyword is to the current keywords, the larger the score is. As σ increases, the range of similarity in the pooling function increases. The kernel features are then fed into a fully connected linear layer to compute a final ranking score:

$$s(\{w_1, \dots, w_h\}, c) = \text{softmax}(W^T \mathcal{K} + b),$$

where $W^{K \times 1}$ and b are trainable parameters. In contrast to the original paper, which uses the tanh activation function, we replace it with the softmax function instead so that the network is learned by minimising the cross entropy between the probabilities and the extracted keywords of the next turn.

Keyword selection Similar to Tang et al.’s model, we take a greedy approach to select the keyword with the highest matching score from the candidate keyword distribution, provided that the keyword is strictly closer to the target than the predicted keyword in the prior turn. The distance between the candidate keyword and target is measured by the inverse of cosine similarity of the corresponding word embeddings.

Conditional keyword response generator Next, we propose to train a separate generative model to generate a contextually relevant response that contains the predicted next keyword \hat{k} . We achieve this by conditioning the keyword in the response generator. During training, the data is manipulated such that the keyword embedding is concatenated to the beginning of the input embedding, separated by a special token [KW], where their position encodings are set to zero. Similar to the above, we fine-tune a GPT-2 decoder to learn the keyword-conditioned generative model:

$$\mathcal{L}_{\text{kernel-gen}} = -\frac{1}{N} \sum_{t=1}^N \log P(r_t | \hat{k}, u_1, \dots, u_n, r_{<t}),$$

where \hat{k} is the embedding of the keyword to be generated and N is the length of the response r .

3.2.3 Persona Generative Model

The persona generative model takes the user persona into account and aims to generate responses that are coherent to the given persona. Since the model focuses on modelling the persona of the user rather than the persona of the agent itself, we also refer to this as the "other-persona generator". To capture the persona information, two methods were considered. First, the persona can be prepended to the utterances, separated by a special token [PER]. Second, the persona can be encoded into a vector representation using a transformer encoder, then decoded in the response generation model. Our experiments found that the concatenation of persona and dialogue context yielded better results. Thus,

we employ this approach and get the following training objective:

$$\mathcal{L}_{\text{persona-gen}} = -\frac{1}{N} \sum_{t=1}^N \log P(r_t | p_1, \dots, p_m \oplus u_1, \dots, u_n, r_{<t}),$$

where N is the length of the response r and \oplus denotes concatenation. Similar to the above, we fine-tune the GPT-2 model, with the role embedding of the persona representing the user.

3.2.4 Persona-Kernel Generative Model

The persona-kernel generative model is the full model aimed to achieve both components. That is, aligning the response with the user persona, and making smooth transitions towards the target. Similar to the kernel generative model, we train a keyword prediction kernel model and a conditional keyword response generation model, but with the following differences.

In addition to the kernel described above, we would like to constrain the keyword candidate to also be close to the persona. Thus, we compute another kernel feature vector with respect to the persona keywords. The context keywords feature vector and the persona keywords feature vector are then combined and fed into a dense layer, which will learn to balance the final score between the context and persona:

$$s(\{w_1, \dots, w_h\}, \{x_1, \dots, x_g\}, c) = \text{softmax}(W'^T(\mathcal{K}_{\text{persona}} \oplus \mathcal{K}_{\text{context}}) + b'),$$

where $\{x_1 \dots, x_g\}$ are keywords of the persona, $\mathcal{K}_{\text{persona}}$ and $\mathcal{K}_{\text{context}}$ are the kernel feature vectors with respect to the persona keywords and context keywords respectively, \oplus represents concatenation, W' and b' are trainable parameters. The next keywords are selected as described above, and a separate conditional keyword response generator is trained to generate the response. Similar to the kernel generative model, the conditional keyword generator is trained by concatenating the keyword embedding in front, separated by the special token [KW]. In addition, the persona is concatenated before the context history, separated by the [PER] special token. Hence, the final training objective is:

$$\mathcal{L}_{\text{persona-kernel-gen}} = -\frac{1}{N} \sum_{t=1}^N \log P(r_t | \hat{k}, p_1, \dots, p_m \oplus u_1, \dots, u_n, r_{<t}),$$

where \hat{k} is the embedding of the next keyword and N is the length of response r .

3.2.5 Summary of models

Table 3.1 summarises the models which be will used in our experiments in Section 4.3. The models can be categorised based on whether they are persona-based or target-guided. In case of persona-based models, both the *other-persona generator* and *persona-kernel generator* incorporate the user’s persona. On the other hand, both the *kernel generator* and the *persona-kernel generator* are target-guided models. The use of the *self-persona generator* will be explained in more detail in Section 4.3.1.

Model	Persona-based	Target-guided
Generator (3.2.1)	No	No
Other-persona Generator (3.2.3)	User’s persona	No
Kernel Generator (3.2.2)	No	Yes
Persona-Kernel Generator (3.2.4)	User’s persona	Yes
Self-persona Generator (3.2.3)	Agent’s persona	No

Table 3.1: Summary of models used in our experiments. The labels within the brackets indicate the corresponding section numbers where detailed explanations of each model can be found. We will compare the first four models as target-guiding agents in our simulation experiment. The self-persona generator is used as a simulated user agent.

Chapter 4

Experiments and Results

4.1 Dataset

Due to the limited availability of real-life datasets containing personas and dialogues of people with dementia, for our research purposes, we solely focus on studying persona-based target-guided agents and turn to an existing dataset **ConvAI2** [14] to train our tasks. Our aim is to evaluate the proposed system to ensure its effectiveness and lay the foundation for future studies in the older adults and dementia population. While these datasets do not specifically capture the nuances of demented persons speech and persona, it serves as a valuable starting point for exploring persona-driven dialogues in our study. In Chapter 5, we will explore the application of our methodology on an older adults dataset and outline a future experimental plan for real-life older adults.

The **ConvAI2** dataset is an extension of the **PersonaChat** dataset, which is created through crowdsourcing via Amazon Mechanical Turk by Zhang et al. [79]. The dataset consists of conversations between randomly paired participants who begin chatting and getting to know each other based on their assigned personas. These personas contain short sentences describing their background, interests, and preferences, such as "I like to ski". By incorporating information from the assigned personas, the dataset aims to address the limitations of traditional chit-chat models and encourages models to generate more consistent and engaging responses. The **ConvAI2** dataset contains a total of 19893 dialogues (17878 training, 1000 validation, 1015 test) of length 4-6 turns each, and 1355 personas (1155 training, 100 validation, 100 test) of 5 sentences each. Table 4.1 shows an example of the **ConvAI2** dataset.

Persona 1	Persona 2
i like to party. my major is business. i'm in college. i love the beach. i work part time at a pizza restaurant.	i enjoy going to museums. i am married. i live in new york city. i have ten nieces and nephews. i'm a writer.
Person 2: hi . how are you doing today ?	
Person 1: hi i am great just finishing up some homework how are you	
Person 2: i'm alright . i just got done writing .	
Person 1: do you write for a living or as a hobby	
Person 2: it is my living . i like culture .	
Person 1: that sounds like a fun job . i am a business major but have a part time job	
Person 2: what are you going to school for	
Person 1: i'm trying to get my ba in finance	
Person 2: do you own your own company	
Person 1: no still in school work at pizza hut part time	
Person 2: do you get free pizza	
Person 1: no but i do get a good deal .	
Person 2: that is a good start	
Person 1: yeah it is . i really hope they have a frat party again soon	
Person 2: i used to party a lot	
Person 1: it is fun i can't get enough	

Table 4.1: Example of ConvAI2 dataset

4.2 PGTask

Ribeiro et al. [57] manipulate the `PersonaChat` dataset to construct `PGDataset` through modelling entailment relations with Natural Language Inference (NLI). The NLI task involves evaluating logical agreement between a pair of sentences, namely entailment, neutral, and contradiction. In the context of `PGDataset`, given an utterance, the authors gather one or more persona sentences that are entailed from that utterance. Table 4.2 shows a few examples of the `PGDataset` and Table 4.3 shows the statistics of the dataset.

Utterance	Persona
yes . two dogs and a cat . they are my babies .	i have two dogs and one cat.
i played a lot of sports in high school , but stopped when i started college .	i am a college graduate. i used to play sports in high school.
haha . i also grew up in a large family on a large farm .	i grew up on a large farm.

Table 4.2: Examples of `PGDataset`

Average number of tokens per response utterance	13.15
Average number of tokens per persona sentence	7.20

Table 4.3: Statistics of `PGDataset`

Implementation details We fine-tune the pretrained `bart-large-cnn` model from Huggingface [74] using PyTorch [50]. We set a batch size of 32 and trained for 3 epochs using one NVIDIA A100 GPU. The AdamW optimiser [41] was used with a learning rate of 2.5e-5 and 3000 warm-up steps. At inference, we generate persona sentences with the fine-tuned model using a beam size of 5, top-p of 0.95 for nucleus sampling [28], and top-k of 3.

Evaluation metrics We evaluate our fine-tuned BART model on the test set of the `PGDataset` and compare its performance against `gpt2-small`, which was the best performing baseline from the original paper. Following their evaluation method, we report the metrics BLEU (precision score which measures the frequency of generated n-grams that are in the ground truth) [49], ROUGE (recall score which measures the number of n-grams in the ground truth that overlap with the generated sentence) [39], and BERTScore (measures similarity based on BERT contextual embeddings) [80] between the generated persona sentences and the ground truth. The results presented in table 4.4 demonstrate that the

fine-tuned BART model outperforms the baseline across all metrics, with significant improvements for longer n-grams (i.e. BLEU-3, BLEU-4, ROUGE-N). This highlights the effectiveness of utilising the encoder-decoder architecture and corrupted inputs in BART compared to the decoder-only GPT-2 model. Through training the model in a denoising approach, the BART model demonstrates an enhanced ability to capture the important persona-related features embedded within utterances, leading to a performance gain in the persona extraction task.

Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	ROUGE-1	ROUGE-2	ROUGE-L	BERTScore
gpt2-small	0.6130	0.3230	0.2062	0.0944	0.5007	0.2831	0.5000	0.9439
bart-large-cnn	0.6477	0.3785	0.2547	0.1807	0.5788	0.3559	0.5656	0.9485

Table 4.4: PGTask results on PGDataset

4.3 Persona-based Target-guided Dialogue Generation

Implementation details For all generative models, we train the pretrained GPT-2 model on the `ConvAI2:original` dataset without revised personas. We fine-tune the `gpt2-medium` model from Huggingface [74], which consists of 24 layers and a hidden size of 1024. Using two NVIDIA A100 GPUs, we train the model with a batch size of 32, learning rate of $2.5e-5$, and the Adam optimiser [32]. At inference, we generate responses using the fine-tuned model with a top-p of 0.95 for nucleus sampling [28], top-k of 3, and maximum response length of 50. For the kernel model, we use 10 RBF kernels with $\mu_k = k/10$ and $\sigma_k = 0.1$. Using one NVIDIA A100 GPU, we train the model with a batch size set to 32, and a learning rate of 0.001 with exponential decay of 0.8 for 16000 steps. For all models, we employ an early stopping strategy. If the validation loss does not decrease for two epochs, training is halted to prevent overfitting.

Evaluation metrics First, we assess the performance of the kernel scoring component in both the kernel generator and persona-kernel generator. Recall that the kernel generator ranks the next keyword candidates using only context keywords, while the persona-kernel generator ranks the candidates using both context and persona keywords. We evaluate the keyword prediction performance for these two ranking kernels on the test set. Specifically, we report the keyword recall at position $K = 1, 3, 5$ over 2677 possible keywords, $R@K$, and the precision at position 1, $P@1$. The results in table 4.5 show that the persona-kernel model has better performance in accuracy, $R@1$, $R@3$, but not $R@5$ compared to the

kernel model. We posit that the persona-kernel model performs better for precision and recall at position 1 because the incorporation of persona keywords enhances the accuracy of predicting the top-ranked keyword. On the other hand, since the kernel model is unconstrained by persona keywords, the model is able to consider a broader range of potential next keywords, leading to a better performance in $R@5$ for the kernel model.

Next, to assess the effectiveness of each model in generating contextually and persona-relevant responses, we evaluate the semantic similarity between the generated responses and both the persona and the context. We compute the cosine similarity of the embeddings of each pair of sentences using SentenceBERT [56]. All persona statements are used to compute the persona similarity and the last two utterances are used to compute the context similarity. Results are shown in table 4.6. The generator, which was trained on context history, serves as the baseline for the other three models. The other-persona generator, which was trained with the persona concatenated to the context, generates responses that most closely align with the user persona, surpassing all three models in persona similarity. On the other hand, the kernel generator achieves the highest degree of context similarity. This is attributed to the greedy approach used in selecting the next keyword, allowing the responses to align more closely with the context. Finally, the persona-kernel generator produces responses that balance both components, ranking second highest in both persona similarity and context similarity among all target-guiding agents. However, it is worth noting that the improvement in performance may be limited due to the nature of the training data. Conversations primarily focus on individuals introducing themselves and discussing their own personas when the speakers try to get to know each other, which may restrict the modelling of the "other" persona.

Model	$R@1$	$R@3$	$R@5$	$P@1$ (Accuracy)
Kernel	0.0545	0.1220	0.1626	0.1119
Persona-kernel	0.0574	0.1231	0.1611	0.1164

Table 4.5: Evaluation of keyword prediction on the ConvAI2 test set

4.3.1 Simulation

In our study, we aim to explore the effects of incorporating persona information into target-guided conversation tasks. Thus, we compare our models with and without persona information. To evaluate the performance of these models in reaching the target, we adopt a simulation-based approach using a similar experimental setup by Tang et al. [65].

Model	Contextual Similarity	Persona Similarity
Generator	0.3941	0.2673
Other-persona Generator	0.4073	0.3177
Kernel Generator	0.4417	0.2670
Persona-Kernel Generator	0.4315	0.3002
Self-persona Generator	0.4050	0.3819

Table 4.6: Evaluation of responses generated by each model. The score is computed using the cosine similarity between each response and context or persona. We compare the first four models for the role of the target-guiding agent. The self-persona generator is used as a simulated user agent.

In each trial, a target keyword and a set of persona statements are randomly selected. A conversation is simulated between a user agent and a target-guiding agent. The user agent plays the role of a human that aligns with the selected persona and is unaware of the target keyword. Meanwhile, the target-guiding agent is given the user’s persona to chat with the user agent, attempting to drive the conversation towards the target.

In order to replicate the user’s responses that maintain a consistent persona, we further train a generative model based on the user’s *self-persona* as opposed to *other-persona*. The self-persona generator model is also trained by concatenating the persona information with the dialogue context, separated by a special token [PER]. The quality of the self-persona agent’s responses is evaluated in Table 4.6, where it is shown to produce responses that have the highest correlation to the persona.

In each trial of the simulation, the target-guiding agent initiates the conversation by randomly selecting a start utterance from a list. We measure the success rate of reaching the target for each agent over 500 trials, as well as the average number of turns required to reach it. A conversation is defined as a success if any utterance contains a keyword which is closely related to the target by a threshold, determined by WordNet’s information content \ln similarity [40]. The threshold for similarity was set to 0.9 in the original paper. However, we find that phrases such as *“that sounds good”* were mistakenly classified as reaching the target *“music”*, even when the response was not directly related. Therefore, we increase the threshold to 0.95. Additionally, we also set a maximum limit of 8 turns for each simulated conversation. If the target is not reached within the limit, the conversation is considered a failure. Thus, when computing the average number of turns to reach the target, only successful cases that reach the target within 8 turns are considered, and failed conversations are excluded.

Table 4.7 shows the results of the simulation experiment for different agents. The

generator model is used as a baseline. The observed drop in success rate when using the self-persona generator as the user agent, in comparison to the generator serving as the user agent, indicates that users who talk more about their own persona tend to influence the target-guiding model to narrow the topics of its responses to the persona, making it not as effective in guiding the conversation towards the target topic.

When chatting with the self-persona generator user agent, using the other-persona generator as a target-guiding agent increases the success rate of reaching the target by 12% compared to using the generator as the target-guiding agent. This shows the benefit of leveraging persona information during the training of the target-guiding agent. Building upon the generator as the baseline for the target-guiding agent when interacting with the self-persona generator, the integration of the kernel generator improves the success rate significantly by 181%, which shows the significance of training a keyword prediction model and conditional keyword response generator in steering the conversation towards the target. Finally, the persona-kernel generator further improves the success rate by another 3%, achieving the highest success rate of 0.680. While the improvement may appear modest, it still highlights the advantages of incorporating persona information into the target-guiding kernel model. Like the performance of persona similarity observed above, we hypothesise training the target-guiding agent with data that emphasises the user’s persona could potentially yield further improvements in the results.

In terms of the average number of turns to reach the target, the results indicate that the generator model has the least number of turns, while the kernel models require more turns on average. However, it is important to note that in the generator model, the target is reached by random, and only successful conversations contribute to this metric. Consequently, if a conversation is successful due to chance rather than the model’s strategy, it will naturally have a lower number of turns. To address this, we run an additional simulation where we increase the maximum number of turns to 50, which will allow us to make a fair comparison of the models’ performances in longer conversations. The results demonstrate that the kernel models require significantly less number of turns to reach the target compared to generator models. In particular, the kernel generator achieves the lowest number of turns, followed by the persona-kernel generator. This shows the benefit of utilising the keyword prediction model to guide the conversation towards the target.

4.3.2 Case Study

In this section, we present a qualitative analysis of the simulated conversations to have a deeper understanding of the strengths and limitations of the models. In order to examine

User agent	Target agent	Success rate (500 trials)	Avg turns (max 8)	Avg turns (max 50)
Generator	Generator	0.254	3.90	14.59
Generator	Kernel Generator	0.684	4.72	6.31
Self-persona Generator	Generator	0.236	3.58	13.64
Self-persona Generator	Other-persona Generator	0.264	3.80	12.10
Self-persona Generator	Kernel Generator	0.662	4.81	5.01
Self-persona Generator	Persona-Kernel Generator	0.680	4.83	7.11

Table 4.7: Simulation results. We compare different target-guiding models where the self-persona generator is used as the user agent.

the differences in incorporating persona information to the target-guiding agent, we compare the examples generated by both the kernel generator and persona-kernel generator. To ensure a fair comparison, we use the same start utterance for each simulated conversation.

Good example In the example shown in Table 4.8, both the persona-kernel generator and kernel generator demonstrate the ability to generate contextually relevant responses and make smooth keyword transitions towards the target *reading*. Notably, the persona-kernel generator is able to introduce a topic related to the user’s persona, like *swimming*, without the user mentioning it first. It is also able to make a connection between the target keyword and the user’s persona even if the topics are not directly related, such as reading books on swimming. This allows the response to relate to the target topic while simultaneously aligning with the user’s interests.

Average example For the example shown in Table 4.9, both the persona-kernel generator and kernel generator successfully make transitions towards the target *dress*. Furthermore, the persona-kernel generator also generates responses that relate to the user’s persona, such as dressing for church. On the other hand, the kernel generator produces responses solely based on the context, as observed by its deviation to the topic of animals. However, both models occasionally generate responses that may sound unnatural. For example, when *hair* is the predicted next keyword, the response *“popular in her hair”* may seem forced.

Failed example Table 4.10 shows an example where both the persona-kernel generator and kernel generator fail to reach the target within the maximum number of turns. Both models struggle to predict the keyword path required to guide the conversation towards

the keyword *exercise*. We investigate that this is due to the first predicted keyword *hobby* already having a cosine similarity of 1.00 with the target keyword, which is the maximum possible value. Consequently, the greedy keyword selection algorithm prevents other keywords from being chosen next, resulting in the target-guiding model being trapped in generating responses around the same keyword *hobby* throughout the conversation. Future work may explore alternative methods to compute similarity between words or more sophisticated keyword selection algorithms.

<p>User persona: i'm a girl who loves sports cars and video games. i only have three best friends. i'm a fourteen year old living with my mom and grandparents. i like swimming.</p>	
<p>Agent: Persona-kernel generator Target: reading</p>	<p>Agent: Kernel generator Target: reading</p>
<p>START: hello how are you tonight ? User agent: i'm great! my best friend is a girl, and she loves video games. how about you? Target agent: <i>[game]</i> i love video games! do you have any hobbies? mine are running and swimming. what about yourself, where is your favorite place to run or swim in the ocean User agent: i am a girl and love swimming. i also like to watch sports cars, especially racing ones! do u? lol what about you guys have any hobbies besides that haha Target agent: <i>[movie]</i> i like to watch movies, but mostly tv. what about you? do u have any hobbies or favorite foods lol User agent: i love sports cars! what do u like to eat? Target agent: <i>[book]</i> i love pizza! what is your favorite book? mine are books on how to swim and swimming, but reading my new one now</p>	<p>START: hello how are you tonight ? User agent: i'm good, just got back from swimming with my best friends. what about you? age 14 living in a home mom and grandparents house now haha but love sports cars too much Target agent: <i>[hobby]</i> i am in high school. my best friend is from france, he's a chef! you? what do u eat for breakfast today? User agent: i'm a girl who likes sports cars and video games. what about you? i am in high school too! age 14 living with mom, grandparents lol hows france cooking up today Target agent: <i>[game]</i> it was nice. i am going to play video games tomorrow! you? what kind of cars do u like best lol nintendo or honda civic for example User agent: i love honda civic and nintendo games. do you have any hobbies? i like to swim, how about yourself what do u play lol haha! Target agent: <i>[write]</i> i like to write, but mostly just play video games lol you? any hobbies or anything else interesting going on in your life today. User agent: i like to play video games, but mostly just watch movies and swim lol what about you? Target agent: <i>[read]</i> i am reading about the storm today lol! what are you doing?</p>

Table 4.8: Good example of simulated conversations guided by the kernel model and persona-kernel model. Predicted next keywords by the target agent are italicised within square brackets.

User persona: i am a 1st grade teacher. i go to church on sunday. my favorite color is purple. i've a 3 year old. my father is a preacher.	
Agent: Persona-kernel generator Target: dress	Agent: Kernel generator Target: dress
START: hi how are you today ?	START: hi how are you today ?
User agent: i am great! my son is 3 and loves to play in the color purple. what are you up too? do u have a church group? we are going to sunday. how about yourself	User agent: i'm good. just got home from church, my favorite sunday! you? what are your hobbies and interests for the week lol
Target agent: <i>[hobby]</i> i am good just watching tv. do u like to cook? my dad taught me when we were little! ery a vegetarian lol what else you enjoy doing for funsies huh	Target agent: <i>[hobby]</i> i'm going to be a vet soon. my favorite animal is the mr gump lol! you sing?
User agent: yes i love cooking! do u have a favorite color? mine is purple. what about you, are there tv shows or movies that interests you at this time? haha my dad was the preacher for awhile now	User agent: i sing in church, my favorite color is purple. you like animals? the only thing we have are our dog and cat
Target agent: <i>[color]</i> i love to cook and have a lot of different colored pasta dishes! what do you like? lol, my favorite color is green. how about yourself eh?	Target agent: <i>[dog]</i> i have a dog, and he loves it! what is your favorite color? mine blue. my dogs name mr gump lol
User agent: my favorite color is purple! i am a teacher, so its hard for me to get out of bed lol. do u like church? we go there often on sundays haha	User agent: i love purple, what is your favorite animal? mine are the dog and cat! you like music then huh
Target agent: <i>[hair]</i> i do not really go to church but my mom is a preacher. she's very popular in her hair! what else? lol, youre married huh? haha how old are they then eh?	Target agent: <i>[girl]</i> yes i do, but my girl is not that into it. you have any pets?
User agent: i am married and have a 3 year old girl! do you have any kids? lol what else u like or dislike about your life	User agent: i've a dog, and she's cute too! you like movies?
Target agent: <i>[dress]</i> no i don't have kids. do you like to dress up for church? what else are u into eh?	Target agent: <i>[dress]</i> i like movies but not as much now that we have a little boy. you dress up for the halloween party?

Table 4.9: Average example of simulated conversations guided by the kernel model and persona-kernel model. Predicted next keywords by the target agent are italicised within square brackets.

<p>User persona: i am a 70 year old man. i live alone in a cabin. i own a boat. i like to go fishing on the lake. i've 5 grandchildren.</p>	
<p>Agent: Persona-kernel generator Target: exercise</p>	<p>Agent: Kernel generator Target: exercise</p>
<p>START: hi how are you today ? User agent: i am good, just relaxing on a cabin deck. you? what do you like to fish in lakes and streams for funsies! haha my grandkids are the best at it too Target agent: <i>[hobby]</i> wow that's awesome i like fishing too. what else do you enjoy? lol! haha, me and my grandkids go to the lake every sunday for fish dinner. where are u from then by chance? User agent: i live alone in a cabin so it is nice to be with my grandkids! i'm not a big fish eater lol haha. what do you have planned for today? we are fishing later Target agent: <i>[hobby]</i> just going to go fishing and eat some fish i love it! how about you? lol. what do u have planned for today User agent: i'm gonna go fishing and have a nice chat with my grandson! he's great. you? lol what do u do for fun in your spare time haha Target agent: <i>[hobby]</i> i like to cook and play with my grandkids lol! what else do you enjoy doing for fun? User agent: i enjoy fishing and cooking! what do you like to cook? haha my grandkids are the best cooks too. it is a pleasure talking with them today Target agent: <i>[hobby]</i> i love to cook and eat fish too! what else do you enjoy fishing for lol?</p>	<p>START: hi how are you today ? User agent: i'm great, just got back from fishing on the lake! you? do anything interesting happen in your life lately that interests ya. mine is my grandkids and i've five of them myself lol what about yours Target agent: <i>[hobby]</i> oh cool. i am in the process of moving to a new place next month and getting married! my fiancée is from france so it will be nice being home soon after she moves here too! User agent: that's cool, i live in the united states but my cabin is located overseas. i'm a 70 year old man and you? Target agent: <i>[hobby]</i> yes i am. we are going to be in houston next friday and you can come by! its so hot there today, any plans this weekend at all? User agent: that's great, i am looking for a boat to go fishing on. i'm not sure if you are in houston but sounds fun! i love fishing irma is coming soon and its so hot there Target agent: <i>[hobby]</i> yes it is. i love to go on kayaks in the summer and hike during fall, its beautiful here this friday! you are awesome lol? do u like music or rap too User agent: i love music! rap is great, but my grandkids are not into that haha. you? do anything interesting happen in your life lately lol Target agent: <i>[hobby]</i> yes i love it! my fiancée is a music major and she's really into the new holland death metal bands. you should come by, they're amazing too haha</p>

Table 4.10: Failed example of simulated conversations guided by the kernel model and persona-kernel model. Predicted next keywords by the target agent are italicised within square brackets.

Chapter 5

Application in Dementia Care

5.1 Experimentation with Older Adults Dataset

We present a new dataset to test the effectiveness of our models in the domain of conversations with older adults. This dataset was collected from a podcast, Senior Storytelling ¹, which was initiated by the Elderly Embrace Care Network, a senior-focused nonprofit organisation. Each episode of the podcast features an interview with a cognitively healthy senior individual, where they share their life stories with a family member, discussing biographical narratives including their childhood, family, interests, and personal values. The list of interview questions used in the podcast is shown in Table 5.1. By experimenting on this dataset, we aim to evaluate our models in a context that align more closely with the conversational dynamics and narratives of older adults. This step is important before proceeding to experiment with real individuals with dementia since the dataset used for training our models initially was not specifically tailored to the older adults population. Testing our models on this new dataset allows us to refine our models that will prepare us for future experiments that involve individuals with dementia.

To create the dataset, four interviews were transcribed into conversational text using the Descript software ². Table 5.2 shows some statistics of our dataset. Since the dataset comprises interview data, the responses are longer in length compared to the training data from the ConvAI2 dataset. We experiment the persona extraction task and persona-based target-guided response generation task on our new dataset.

¹<https://podcasts.google.com/feed/aHR0cHM6Ly9hbmNob3IuZm0vYcy8yODQ2MmWYwNC9wb2RjYXNOL3Jzcw?sa=X&ved=0CBUQ27cFahcKEwj468GEpIf9AhUAAAAAHQAAAAAQaw>

²<https://www.descript.com/transcription>

Question
1 Would you like to introduce yourself?
2 Who was your biggest role model as a young person and why?
3 If you could relive one year or decade of your life, what would it be and why?
4 If you could speak to your younger self, what advice would you give?
5 Do you have any regrets from your younger years that you would feel is a good lesson for a younger generation?
6 Are there any habits or skills that you regret taking, picking, or not picking?
7 What advice did you get that you wish you would've listened to or taken more serious?
8 What is something about today's youth that really surprises you or is fundamentally different from when you were our age?
9 Is there something about today's world that you never expected to become the way that it is?
10 What is something that was common in your youth that is now obsolete?
11 Nowadays we have companies like Amazon, Uber and things like that that make our lives easier. So what was something that was like those companies in your day?
12 What's something that you really enjoy doing something that brings you a lot of joy in your life?
13 Do you have a go-to comfort food and how has that changed throughout your life?
14 Do you have any funny stories that were legal back in your day, but not anymore?
15 What are some of the little things that you do to relax or take your mind off things?
16 What is your biggest milestone in your life to date?
17 What is your favorite childhood memory?
18 What career would you have pursued if you, if there wasn't any concern with money or practicality?
19 How did you and your spouse meet? And what do you remember most about your wedding?
20 What were your fondest memories of school or education in general? And how do you think it compares today?
21 What are some life lessons that your parents taught you that are no longer applicable?
22 Do you have any advice or wisdom for kids today considering how much technology affects relationships and interactions?
23 After seeing how much the world has changed until now. Do you have any predictions of what it will look like in the upcoming future?
24 What was the hardest and or easiest part of getting older?
25 When you were young, what did you think 2020 would look like? What sort of advances in technology did you think that we would have?
26 If you have any story from your childhood that you'd like people to know now, What would it be?

Table 5.1: Interview questions of our older adults dataset

Average number of turns	30.0
Average number of tokens per response	127.5

Table 5.2: Statistics of our older adults dataset

5.1.1 Persona Extraction

We first extract personas from our older adults dataset. Tables 5.3 and 5.4 show the first four persona statements generated by the fine-tuned BART and GPT-2 models from the interview transcript of two older adults, Mrs. R and Mr. M, respectively. As observed in the results, the GPT-2 baseline is not able to generate complete sentences from the utterances. We suspect this is due to training with shorter input lengths in the `PGDataset` data (average 7.2 tokens). Consequently, when faced with complex and longer sentence inputs in our older adults dataset (average 127.5 tokens), GPT-2 struggles to extract the most important features of the full context, leading to difficulties in generating the corresponding persona sentences.

In contrast, the bidirectional encoder and integration of corrupted inputs in BART allow the model to be robust to longer sequences, enabling it to learn richer representations of the inputs and generate coherent persona sentences from the response utterances. Overall, BART produces concise personas that accurately reflect the corresponding responses of both Mrs. R and Mr. M. Like the `PGDataset`, the extracted personas describe the individual’s location, family, history, hobbies, and values. For example, the fine-tuned BART model inferred from Mrs. R’s responses that she lives in Santa Clara, has six daughters and four sons, and grew up in a valley, whereas Mr. M enjoys reading and making friends. However, there are a few exceptions where the generated personas are inconsistent with the responses, which are discussed in more detail in Section 5.3. We suppose that conducting an interview of similar structure with demented people who are in early stages will be useful for collecting their personas for future applications.

5.1.2 Persona-based Target-guided Response Generation

Following the persona extraction step, we simulate conversations between an imitated user from our older adults dataset and a persona-based target-guiding agent. Similar to the setup in Section 4.3.1, the responses of the imitated user are generated by the self-persona generator based on the inferred personas from the interview transcript. Considering the potential application in the field of dementia care, we experiment on the target topics "eat" and "bath" in our simulations. Additionally, a common caregiver practice is to establish

Interview transcript	Extracted persona (BART)	Extracted persona (GPT-2)
<p>Interviewer: Before we start do you want to say a little bit about yourself?</p> <p>Mrs. R: First of all, I still have a strong voice at the age of 90, so a lot of people on the phone don't think I'm 90, but this year I am 90 and I have lived in Santa Clara.</p>	i live in Santa Clara.	i love.
<p>Interviewer: If you could relive one year or decade in of your life, what would it be and why?</p> <p>Mrs. R: ... I had a happy life so, I really don't know that I'd wanna live any of my years again. They were, I had wonderful times raising our children. We decided to have a large family and so we ended up with six daughters and four sons and they're all alive and doing well today. Very proud of them ...</p>	i have six daughters and four sons.	i am apart.
<p>Interviewer: Thanks grandma. If you could speak to your younger self, what advice would you give?</p> <p>Mrs. R: If you could speak to your younger self, what advice would you give? Hmm. Well, I don't, it was really interesting ... I went to college for one year, and then I was a mom and I really loved being a mom. So we had numerous children and they were all very active doing things, making things, being active. Boys all had paper routes and had a little bit of money that way, and the girls did a lot of babysitting and a lot of sewing, and a lot of helping canning fruit. We live in a valley that had lots and lots of wonderful fruit. Apricots, pears, plumbs prunes. So everybody was kind of involved with the fruit industry. It's a wonderful place to grow up.</p>	i grew up in a valley. The fruit industry was big in the area.	.
<p>Interviewer: That's nice. Thank you. Mm-hmm. , was there anything that you wanted to do but never did and why did you Never do said things?</p> <p>Mrs. R: ... And in those days anyway, there was a lot of different things that we did with our children and they learned all kinds of things. They, I had them Involved in traveling and all kinds of things and camping and just doing all kinds of things, going to the beach. And so they, they traveled. They weren't afraid of their world. It was my intention that my children would all be not afraid of the world. My father taught me not to be afraid of the world and enjoy it. And that's what I try to impart with our children.</p>	my father taught me not to be afraid of the world.	.

Table 5.3: Examples of extracted personas in an interview transcript with Mrs. R

Interview transcript	Extracted persona (BART)	Extracted persona (GPT-2)
<p>Interviewer: Just to start. Mm-hmm. , uh, when you were younger, who is your biggest role model and why?</p> <p>Mr. M: My model is actually is, uh, my, uh, classmate, uh, chow Wong. And, uh, you know why it is, uh, because, uh, he is good in the, the writing skill. And when I was young, my writing skill was not So good cuz his writing skill is so good. So he was my model.</p>	my writing is very good.	you know why it is, because, ugh, chow wong is my favorite author.
<p>Interviewer: Okay. If you could relive one year or decade or any portion of your life, uh, what would it be and why?</p> <p>Mr. M: ... I can do a lot of things and the extra things is, uh, beyond my study, so I go to exploring. A lot of meaningful things in my life. However, if I do this way, I will lose my, a lot of my opportunities cuz uh, and I will lose my progress in the, in this one year because all the progress is accumulation. So if I do. So I will lose my progress, but in other ways, uh, if I, uh, believe this one year I may be, could exploring other meaningful, uh, than my developing my interests, uh, like a sports. Like hobbies. That's it.</p>	i am exploring a lot of things beyond my studies.	i love to read a lot of books.
<p>Interviewer: Okay. Hey, um, do you have any regrets from your younger years that you feel would be a good lesson to younger generations?</p> <p>Mr. M: Yes. My advice to myself is to read more books. I think that. I was in lack of reading. Uh, more books, reading books, giving me a lot of ideas, all I could, accumulating my knowledges and this, which is beyond my class. A lot of things I won't be able to learn in my class, but I can learn through the reading books and this ...</p>	i love to read.	.
<p>Interviewer: Was there anything that you really wanted to do in your life but never did? And why did you never do that thing?</p> <p>Mr. M: ... during my study in the young ages, I was very much a focus to studying er the subjects in class, cuz I lose a. Since, uh, to exploring other things beyond my class. One is, uh, uh, reading our books to discover other things, uh, variety of books are developing. My other interests, uh, these interests including are, could be a sports, uh, making a friends, especially the friends I like, uh, to make. During the young time, especially in the, in the middle school or high school, you're making friends, close friends, including, uh, boyfriends or girlfriends, but everything you know, is more enjoyable ...</p>	i enjoy making friends.	.

Table 5.4: Examples of extracted personas in an interview transcript with Mr. M

a connection to the individual’s persona and build comfort around the target topic [68]. Thus, we also investigate whether user interest is maintained after reaching the target in a conversation. Specifically, we allow the simulation to continue for two turns after the target is reached and examine whether the topic sustains. Tables 5.5 and 5.6 show successful examples of reaching the targets in simulated conversations for Mrs. R and Mr. M respectively.

For the target *eat*, the persona-kernel generator is able to connect Mrs. R’s persona *soup dinner* with the target immediately. Similarly, Mr. M’s persona of *reading* is integrated into the target agent’s response while transitioning to the *food* topic. Both simulated user agents Mrs. R and Mr. M are also able to carry on the conversation around the subject of food by expanding their discussions about *pasta* and *meatlovers* respectively. Furthermore, the persona-kernel generator tailors its responses to align with the user’s persona. For example, in Mrs. R’s simulation, the target agent mentions not liking Facebook, and in Mr. M’s simulation, the target agent expresses a liking for coffee.

The target *bath* is a more challenging problem since it has no apparent relationship with the personas of either Mrs. R or Mr. M. Nevertheless, the persona-kernel generator adeptly selects transition keywords in each turn that lead to the target. In Mrs. R’s simulation, the target agent is able to respond according to the context of previous turns, such as asking about the user’s situation in *california* and the *cold weather*. In general, the transitions are smooth and relevant as the target agent incorporates the target *bath* into its responses by talking about *warm water baths*. However, there are relatively fewer mentions of Mrs. R’s persona in the target agent’s responses, possibly due to its focus on reaching the target within the given context. In Mr. M’s simulation, there are more discussions of Mr. M’s persona in the target agent’s responses, including *reading*, *coffee*, and *helping people in need*. After the target agent successfully reaches the target, the simulated user agent diverges from the target topic and talks about its own interest in *coffee*. However, the target agent continues the conversation by bridging the topic of *coffee* back to the target keyword *bath*, prompting the simulated user agent to respond about *baths* as well. This demonstrates the potential of sustaining the user’s interest in the target topic even after it has been achieved. Nevertheless, the responses of the target agent occasionally feel slightly forced as it attempts to incorporate the predicted keyword. For example, the phrase “*read about people that help others with their baths*” sounds unnatural despite mentioning the persona and target, indicating room for improvement in generating more fluent and appropriate responses.

Extracted persona of Mrs. R

i live in Santa Clara.

i am not a big fan of Facebook.

i worry too much. whenever i worry it's time to make soup.

i enjoy cooking and baking.

Agent: Persona-kernel generator **Target:** eat

Target agent: hi . how are you doing today ?

Mrs. R: i'm great! what are you up to? facebook makes me anxious. it scares the heck out of my husband and i! do not like that. how about a soup dinner?

Target agent: *[eat]* i'm not good at facebook. my husband hates it, though! lol what do you like to eat? mine is pasta and meatloaf

Mrs. R: i love pasta! do you like to bake? i'm not good at baking but enjoy making soup. what about foodie hobbies eh

Target agent: *[eat]* i like eating meatloaf! do you cook for your family? my husband is a chef

Mrs. R: i cook for my family. what about you? do you like meatloaf or pasta!

Target agent: *[eat]* i like eating meatloaf.

Agent: Persona-kernel generator **Target:** bath

Target agent: hi how are you doing today

Mrs. R: i am good. just got done making soup! and you? what do you cook and bake for friends

Target agent: *[hobby]* i love it here in london, where my mom lives. you? lol what else do u like to cook and bake for friends

Mrs. R: well, i do not like facebook either! irma is a horrible time of year for us here in california.

Target agent: *[house]* yes i am a bit of an outdoorsy kinda gal. my house is on the edge and its pretty much bare ground now lol irmas are horrible rightnow, so sorry about that

Mrs. R: well, it was nice to chat with you! do not worry about me, i am just relaxing and making soup! so what are the hobbies of yours that interests you

Target agent: *[water]* i love the water. its so peaceful and relaxing! how do you like it in california? lol

Mrs. R: well, i'm a little sad about the weather. its just not too warm here in california!

Target agent: *[bath]* it gets cold here in london. i prefer the colder weather, but my house is on a lot of empty land so its hard to stay warm!

Mrs. R: i bet its a lot colder there! lol. do you like to cook? my favorite is making soups and baking cookies

Target agent: *[bath]* yes! i like to make a lot of them, especially for my bath and cooking classes. how about you?

Mrs. R: i love making soups and cookies! it is a lot of fun! so what else you like to do? lol

Target agent: *[bath]* i love going to the bath! its so relaxing. how about you? london is great here in springtime too, but summer weather can get a little cold

Mrs. R: i like the cold weather too! lol. do you have any hobbies besides baking cookies?

Target agent: *[bath]* yes! i love the warm water in my tubs and baths. how many times have you made a soup?

Table 5.5: Examples of simulated conversations between Mrs. R and the target agent

Extracted persona of Mr. M

i love to read.
i went to the university of aeronautics and astronautics.
my father always told me to help people in need.
coffee is important to me.

Agent: Persona-kernel generator **Target:** eat

Target agent: hi how are you today
Mr. M: i'm good just finished reading a book about coffee. any hobbies?
Target agent: [*hobby*] i love to read and listen too. do you like coffee? my favorite is cappuccino! lol, what are your hobbies though
Mr. M: i like coffee, but not as much now that it is hard to find a good cappuccino in the market. what do you read?
Target agent: [*food*] i like to read about the world of food. do you have any hobbies? mine is cooking, lol!
Mr. M: i like to read about the world of food and travel. it has been a rough year, what is it?
Target agent: [*eat*] i am glad it is over. what else do you like to eat? my favorite food are meat lovers and coffee fiends though lol!
Mr. M: meat lovers are the best! i like to eat chicken and fish
Target agent: [*eat*] i like to cook and eat meat too. do you have any pets? what about coffee is your favorite thing!
Mr. M: yes, i love coffee! what kind of food are you into? do u have any pets or hobbies besides cooking
Target agent: [*eat*] i love to eat meat, but not as much coffee haha lol! how about yourself?

Agent: Persona-kernel generator **Target:** bath

Target agent: good afternoon how are you today
Mr. M: i am good, just reading a book. what do you read? my dad always said to help people in need! he is dead now but taught me how important it was for him and his family!
Target agent: [*hobby*] i love books too! what else do you read? mine is history. iced tea helps me sleep though
Mr. M: my dad was a space pilot! he taught me about the importance of helping others, especially those less fortunate than myself
Target agent: [*water*] my parents are still alive though and we live near the ocean so its safe for me to swim! do u like water sports at all or only surfers
Mr. M: do you swim? my dad taught me to help people, especially those less fortunate than myself
Target agent: [*bath*] what else u read about people that help others with their baths or baths and showers
Mr. M: i read about people that help those in need, especially my dad and he taught me how important coffee is to sleep.
Target agent: [*bath*] coffee is important to me as well, i drink a ton of it while in the bath. do you like water baths too! haha what kind of things have been done for help people that need them though?
Mr. M: i do not like water baths! lol what is your favorite book? mine was about people that help others. italian novels are my fave
Target agent: [*bath*] i like water baths too! what do u think of them though lol, haha.

Table 5.6: Examples of simulated conversations between Mr. M and the target agent

5.2 Moving Towards Dementia: Experimentation Plan for Older Adults

Recognising the challenges associated with technology engagement and empirical data-gathering with people with dementia, we acknowledge the importance of first conducting experiments with healthy older adults. This initial step will help address any accessibility issues specific to this population, ensuring a solid foundation before proceeding to experiments with individuals with dementia, who may face additional complexities to do with consent, language abilities, and challenging behaviours. The following plan is a rough sketch of what such an initial experiment involving non-demented older adults might look like. This preliminary stage will allow refinement of the methodology such that a future experiment could be adapted to individuals with dementia.

- **Participant recruitment:** Obtain ethical approval and recruit participants from local older adult populations, screening to exclude potential dementia diagnoses. Gather demographic information such as gender, cultural background, and familiarity with technology.
- **Persona collection:** Collect a diverse set of narrative dialogues from participating older adults. This could be in a life story interview format similar to the podcast with older adults described above. Next, adapt the fine-tuned BART model on the collected data to extract personas for each individual, then have participants evaluate the correctness of the extracted personas and make corrections where applicable.
- **Experimental setup:** Initiate a user study that involves interactions between each participant and the conversational agent conditioned on their persona. The target of the conversation may be set to topics related to daily activities, such as bath, eat, medication, etc. similar to the simulated conversations in the above section. Text-to-speech or speech-to-text technologies may be employed on the interface depending on individual needs.
- **Evaluation metrics:** In addition to measuring the success rate and number of turns needed to reach the target topic, participants will complete a survey before and at the end of the interaction stating their interest in the target topic, user satisfaction, the smoothness of transitions, and alignment of the generated responses with their identities and preferences. These metrics can be rated on a scale 1-10. The conversation data will also be recorded for evaluation from other humans and caregivers.

- **Iterative improvement:** Incorporate the insights gathered from the participants to refine the methodology. The models may also be fine-tuned on the newly collected data in order to capture the linguistic characteristics of older adults. Annotators can label new persona statements from the interview dialogues and real-life conversations with caregivers may be injected into the data. We expect the collection of user persona focused conversations will improve the performance of the persona-kernel generator model.

5.3 Limitations

Model inaccuracies Both the persona extraction and response generation steps rely on generative models, which may occasionally hallucinate and generate inaccurate information. For example, the inferred personas sometimes contradict with the ground truth (Table 5.7). Furthermore, due to the nature of the training data, which contains utterances of shorter lengths, the fine-tuned model may not capture all persona information when a response exceeds a certain length during the inference stage, as observed in the interview dialogue data. For instance, Mrs. R’s age is not extracted as a persona sentence in Table 5.3.

Utterance	Extracted persona
My model is actually is, uh, my, uh, classmate, uh, chow Wong. And, uh, you know why it is, uh, because, uh, he is good in the, the writing skill. And when I was young, my writing skill was not So good cuz his writing skill is so good. So he was my model.	my writing is very good.

Table 5.7: Erroneous example of persona generation

Dementia language Training and evaluation of our models rely solely on publicly available data, which does not represent the population of people with dementia. Dementia can affect the language abilities of the individual, such as word-finding difficulty, grammatical deficits, and a lack of cohesion in discourse [31]. The variations in language use may pose challenges for the conversational agent in understanding and generating appropriate responses.

Potential risks As with other personalised systems, the collection of user personas may raise concerns about privacy and ethics. There is a risk for personal information being carelessly shared with other users or exploited for manipulative purposes, such as targeted advertising, which can cause harm to vulnerable populations like older adults [26]. In addition, it is important to consider potential biases in conversational models, particularly given the unknown demographics and cultural diversity of **ConvAI2**. Unintended biases in the training data can unwittingly influence the responses generated, possibly leading to exclusion of certain groups of people. For example, conversational agents may fail to adapt to gender, racial, or cultural disparities when selecting topics to chat about. Therefore, it is necessary to train the model on a diverse dataset that encompasses individuals from various cultures and with different language abilities. Constant human supervision is essential to identify and mitigate these risks.

Chapter 6

Conclusion

In this thesis, we proposed a framework for building a persona-based target-guided dialogue system that holds potential for application in the domain of dementia care. Our methodology involves a persona extraction step and a persona-based target-guided conversational agent which interacts with the user. We experimented our models on a general public dataset ConvAI2. In the persona extraction task, we found that by fine-tuning BART, the model showed improved robustness to noise, outperforming the GPT-2 baseline on the PGTask evaluation. Next, we trained a persona-based keyword transition kernel and a conditional keyword generative model for the conversational agent. This model aims to generate contextually relevant responses that not only align with the user’s persona but also guide a conversation towards a predefined target topic. By integrating responses that align with the user’s persona, this approach aims to foster engagement and arouse user interest in the target topic. We tested the models using a self-simulation approach and found that the incorporation of persona information and keyword transition kernel increased the success rate of reaching the target.

To further investigate the applicability of our models within the older adults population, we tested both components of this framework on our newly constructed older adults dataset. Our models are able to generate mostly accurate personas from dialogue data and carry out personalised conversations that successfully reach the set targets.

To the best of our knowledge, this is the first work studying conversational AI models designed for the domain of dementia care. We believe the proposed framework can address the scarcity issue of caregivers and provide an accessible, personalised tool to assist individuals with dementia. Although it has not been directly tested on people with dementia due to limited resources, our preliminary experimentation with a real-life dataset

of older adults' narratives shows notable results, which encourages further development in this area.

6.1 Future Work

Dynamic updates For future work, persona information can potentially be updated dynamically through interacting with the conversational agent. The real-time updates of the model allow responses to be more relevant to the current state of the user, which can enhance engagement.

Emotional aspect Conversations offer a rich source of emotional signals. Many existing persona-based models do not consider the sentiment aspect. For example, a persona statement like "I have two sons" does not capture the user's attitudes towards their children. Incorporating the emotions into the model can help generate responses that align with the user's expectations. Additionally, the agent may be more empathetic by adapting its tone to be more passive or assertive based on the user's mood or portrayed identity.

Experimentation with older adults Following the proposed plan in Section 5.2, future work may conduct real-life experiments with older adults to evaluate the effectiveness of the system. The data and feedback collected from their interactions can be used to fine-tune the models for this domain, so that the system can better cater to the needs and communication patterns of older adults.

References

- [1] Alzheimer’s Association. 2023 alzheimer’s disease facts and figures. *Alzheimer’s & Dementia*, 19(4):1598–1695, 2023.
- [2] Rajat Kumar Behera, Pradip Kumar Bala, and Arghya Ray. Cognitive chatbot for personalised contextual customer service: Behind the scene and beyond the hype. *Information Systems Frontiers*, pages 1–21, 2021.
- [3] David A Bennett, Julie A Schneider, Yuxiao Tang, Steven E Arnold, and Robert S Wilson. The effect of social networks on the relation between alzheimer’s disease pathology and level of cognitive function in old people: a longitudinal cohort study. *The Lancet Neurology*, 5(5):406–412, 2006.
- [4] Steven Bird and Edward Loper. NLTK: The natural language toolkit. In *Proceedings of the ACL Interactive Poster and Demonstration Sessions*, pages 214–217, Barcelona, Spain, July 2004. Association for Computational Linguistics.
- [5] Laura Block, Andrea Gilmore-Bykovskiy, Anna Jolliff, Shannon Mullen, and Nicole E Werner. Exploring dementia family caregivers’ everyday use and appraisal of technological supports. *Geriatric Nursing*, 41(6):909–915, 2020.
- [6] Robert N Butler. The life review: An interpretation of reminiscence in the aged. *Psychiatry*, 26(1):65–76, 1963.
- [7] Guanyi Chen, Yinhe Zheng, and Yupei Du. Listener’s social identity matters in personalised response generation. *arXiv preprint arXiv:2010.14342*, 2020.
- [8] Ruijun Chen, Jin Wang, Liang-Chih Yu, and Xuejie Zhang. Learning to memorize entailment and discourse relations for persona-consistent dialogues. *arXiv preprint arXiv:2301.04871*, 2023.

- [9] Jiale Cheng, Sahand Sabour, Hao Sun, Zhuang Chen, and Minlie Huang. Pal: Persona-augmented emotional support conversation generation. *arXiv preprint arXiv:2212.09235*, 2022.
- [10] Sheung-Tak Cheng. Dementia caregiver burden: a research update and critical analysis. *Current psychiatry reports*, 19:1–8, 2017.
- [11] Itsugun Cho, Dongyang Wang, Ryota Takahashi, and Hiroaki Saito. A personalized dialogue generator with implicit user persona detection. *arXiv preprint arXiv:2204.07372*, 2022.
- [12] Kenneth Mark Colby. *Artificial paranoia: A computer simulation of paranoid processes*, volume 49. Elsevier, 2013.
- [13] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [14] Emily Dinan, Varvara Logacheva, Valentin Malykh, Alexander Miller, Kurt Shuster, Jack Urbanek, Douwe Kiela, Arthur Szlam, Iulian Serban, Ryan Lowe, et al. The second conversational intelligence challenge (convai2). In *The NeurIPS'18 Competition: From Machine Learning to Intelligent Conversations*, pages 187–208. Springer, 2020.
- [15] Wenhan Dong, Shi Feng, Daling Wang, and Yifei Zhang. I know you better: User profile aware personalized dialogue generation. In *International Conference on Advanced Data Mining and Applications*, pages 192–205. Springer, 2022.
- [16] Joseph Enright et al. *Identity and Caregiver Burden in Dementia: An Evaluation of Telehealth Reminiscence for Informal Caregivers*. PhD thesis, University of Saskatchewan, 2016.
- [17] Heather Eustace. Why i want to turn my oma into a youtuber. Masters Research Project, University of Waterloo, 2023.
- [18] Sam Fazio, Douglas Pace, Janice Flinner, and Beth Kallmyer. The fundamentals of person-centered care for individuals with dementia. *The Gerontologist*, 58(suppl.1):S10–S19, 2018.
- [19] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health*, 4(2):e7785, 2017.

- [20] Joana Galvão Gomes da Silva, David J Kavanagh, Tony Belpaeme, Lloyd Taylor, Konna Beeson, and Jackie Andrade. Experiences of a motivational interview delivered by a robot: qualitative study. *Journal of medical Internet research*, 20(5):e116, 2018.
- [21] Moojan Ghafurian, Jesse Hoey, and Kerstin Dautenhahn. Social robots for the care of persons with dementia: a systematic review. *ACM Transactions on Human-Robot Interaction (THRI)*, 10(4):1–31, 2021.
- [22] Gemma Goodall, Lara André, Kristin Taraldsen, and J Artur Serrano. Supporting identity and relationships amongst people with dementia through the use of technology: A qualitative interview study. *International Journal of Qualitative Studies on Health and Well-being*, 16(1):1920349, 2021.
- [23] Jia-Chen Gu, Zhen-Hua Ling, Yu Wu, Quan Liu, Zhigang Chen, and Xiaodan Zhu. Detecting speaker personas from conversational texts. *arXiv preprint arXiv:2109.01330*, 2021.
- [24] Jia-Chen Gu, Hui Liu, Zhen-Hua Ling, Quan Liu, Zhigang Chen, and Xiaodan Zhu. Partner matters! an empirical study on fusing personas for personalized response selection in retrieval-based chatbots. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 565–574, 2021.
- [25] Prakhar Gupta, Harsh Jhamtani, and Jeffrey P Bigam. Target-guided dialogue response generation using commonsense and data augmentation. *arXiv preprint arXiv:2205.09314*, 2022.
- [26] Woodrow Hartzog. Unfair and deceptive robots. *Md. L. Rev.*, 74:785, 2014.
- [27] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- [28] Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*, 2019.
- [29] William S Horton, Daniel H Spieler, and Elizabeth Shriberg. A corpus analysis of patterns of age-related change in conversational speech. *Psychology and aging*, 25(3):708, 2010.
- [30] Hongyan Jing, Nanda Kambhatla, and Salim Roukos. Extracting social networks and biographical facts from conversational speech transcripts. In *Proceedings of the 45th*

- Annual Meeting of the Association of Computational Linguistics*, pages 1040–1047, 2007.
- [31] Daniel Kempler and Mira Goral. Language and dementia: Neuropsychological aspects. *Annual review of applied linguistics*, 28:73–90, 2008.
- [32] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [33] Yosuke Kishinami, Reina Akama, Shiki Sato, Ryoko Tokuhisa, Jun Suzuki, and Kentaro Inui. Target-guided open-domain conversation planning. *arXiv preprint arXiv:2209.09746*, 2022.
- [34] Tom Kitwood. The experience of dementia. *Aging & mental health*, 1(1):13–22, 1997.
- [35] Alyssa Kubota, Emma IC Peterson, Vaishali Rajendren, Hadas Kress-Gazit, and Laurel D Riek. Jessie: Synthesizing social robot behaviors for personalized neurorehabilitation and beyond. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*, pages 121–130, 2020.
- [36] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*, 2019.
- [37] Jiwei Li, Michel Galley, Chris Brockett, Georgios P Spithourakis, Jianfeng Gao, and Bill Dolan. A persona-based neural conversation model. *arXiv preprint arXiv:1603.06155*, 2016.
- [38] Shuokai Li, Ruobing Xie, Yongchun Zhu, Xiang Ao, Fuzhen Zhuang, and Qing He. User-centric conversational recommendation with multi-aspect user modeling. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 223–233, 2022.
- [39] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81, 2004.
- [40] Dekang Lin et al. An information-theoretic definition of similarity. In *Icml*, volume 98, pages 296–304, 1998.
- [41] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

- [42] Hongyuan Lu, Wai Lam, Hong Cheng, and Helen Meng. Partner personas generation for dialogue response generation. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5200–5212, 2022.
- [43] Carol Ann Maher, Courtney Rose Davis, Rachel Grace Curtis, Camille Elizabeth Short, and Karen Joy Murphy. A physical activity and diet program delivered by artificially intelligent virtual health coach: proof-of-concept study. *JMIR mHealth and uHealth*, 8(7):e17558, 2020.
- [44] Seyed Mahed Mousavi, Alessandra Cervone, Morena Danieli, and Giuseppe Riccardi. Would you like to tell me more? generating a corpus of psychotherapy dialogues. In *Proceedings of the Second Workshop on Natural Language Processing for Medical Conversations*, pages 1–9, 2021.
- [45] Seyed Mahed Mousavi, Roberto Negro, Giuseppe Riccardi, et al. An unsupervised approach to extract life-events from personal narratives in the mental health domain. In *CEUR WORKSHOP PROCEEDINGS*, volume 3033, pages 1–5. Associazione Italiana di Linguistica Computazionale, 2021.
- [46] Ursula Naue and Thilo Kroll. ‘the demented other’: identity and difference in dementia. *Nursing Philosophy*, 10(1):26–33, 2009.
- [47] World Health Organization. Dementia fact sheet, 2023.
- [48] Serguei VS Pakhomov, Eden A Kaiser, Daniel L Boley, Susan E Marino, David S Knopman, and Angela K Birnbaum. Effects of age and dementia on temporal cycles in spontaneous speech fluency. *Journal of neurolinguistics*, 24(6):619–635, 2011.
- [49] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318, 2002.
- [50] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [51] Olimpia Pino, Giuseppe Palestra, Rosalinda Trevino, and Berardina De Carolis. The humanoid robot nao as trainer in a memory program for elderly people with mild cognitive impairment. *International Journal of Social Robotics*, 12:21–33, 2020.

- [52] Chloé Pou-Prom, Stefania Raimondo, and Frank Rudzicz. A conversational robot for older adults with alzheimer’s disease. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(3):1–25, 2020.
- [53] Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. Improving language understanding by generative pre-training. 2018.
- [54] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019.
- [55] Delip Rao, David Yarowsky, Abhishek Shreevats, and Manaswi Gupta. Classifying latent user attributes in twitter. In *Proceedings of the 2nd international workshop on Search and mining user-generated contents*, pages 37–44, 2010.
- [56] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019.
- [57] Rui Ribeiro, Joao P Carvalho, and Luísa Coheur. Pgtask: Introducing the task of profile generation from dialogues. *arXiv preprint arXiv:2304.06634*, 2023.
- [58] Steve R Sabat and Rom Harré. The construction and deconstruction of self in alzheimer’s disease. *Ageing & Society*, 12(4):443–461, 1992.
- [59] Steven R Sabat. Epistemological issues in the study of insight in people with alzheimer’s disease. *Dementia*, 1(3):279–293, 2002.
- [60] Karin Sevegnani, David M Howcroft, Ioannis Konstas, and Verena Rieser. Otters: One-turn topic transitions for open-domain dialogue. *arXiv preprint arXiv:2105.13710*, 2021.
- [61] Marianne Smith and Kathleen Buckwalter. Behaviors associated with dementia: Whether resisting care or exhibiting apathy, an older adult with dementia is attempting communication. nurses and other caregivers must learn to ‘hear’this language. *AJN The American Journal of Nursing*, 105(7):40–52, 2005.
- [62] Yueming Sun and Yi Zhang. Conversational recommender system. In *The 41st international acm sigir conference on research & development in information retrieval*, pages 235–244, 2018.

- [63] Claire Alice Surr. Preservation of self in people with dementia living in residential care: A socio-biographical approach. *Social science & medicine*, 62(7):1720–1730, 2006.
- [64] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27, 2014.
- [65] Jianheng Tang, Tiancheng Zhao, Chenyan Xiong, Xiaodan Liang, Eric P Xing, and Zhiting Hu. Target-guided open-domain conversation. *arXiv preprint arXiv:1905.11553*, 2019.
- [66] Anna Tiginova, Andrew Yates, Paramita Mirza, and Gerhard Weikum. Charm: Inferring personal attributes from conversations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5391–5404, 2020.
- [67] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [68] Aline Vézina, Line Robichaud, Philippe Voyer, and Daniel Pelletier. Identity cues and dementia in nursing home intervention. *Work*, 40(1):5–14, 2011.
- [69] Kazuyoshi Wada and Takanori Shibata. Living with seal robots in a care house—evaluations of social and physiological influences. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4940–4945. IEEE, 2006.
- [70] Kazuyoshi Wada, Takanori Shibata, Tomoko Saito, and Kazuo Tanie. Effects of robot-assisted activity for elderly people and nurses at a day service center. *Proceedings of the IEEE*, 92(11):1780–1788, 2004.
- [71] Jian Wang, Dongding Lin, and Wenjie Li. Follow me: Conversation planning for target-driven recommendation dialogue systems. *arXiv preprint arXiv:2208.03516*, 2022.
- [72] Xuwei Wang, Weiyan Shi, Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, and Zhou Yu. Persuasion for good: Towards a personalized persuasive dialogue system for social good. *arXiv preprint arXiv:1906.06725*, 2019.
- [73] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Commun. ACM*, 9(1):36–45, jan 1966.

- [74] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Huggingface’s transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*, 2019.
- [75] Chien-Sheng Wu, Andrea Madotto, Zhaojiang Lin, Peng Xu, and Pascale Fung. Getting to know you: User attribute extraction from dialogues. *arXiv preprint arXiv:1908.04621*, 2019.
- [76] Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. Proactive human-machine conversation with explicit conversation goals. *arXiv preprint arXiv:1906.05572*, 2019.
- [77] Chenyan Xiong, Zhuyun Dai, Jamie Callan, Zhiyuan Liu, and Russell Power. End-to-end neural ad-hoc ranking with kernel pooling. In *Proceedings of the 40th International ACM SIGIR conference on research and development in information retrieval*, pages 55–64, 2017.
- [78] An-Zi Yen, Hen-Hsen Huang, and Hsin-Hsi Chen. Personal knowledge base construction from text-based lifelogs. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 185–194, 2019.
- [79] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. Personalizing dialogue agents: I have a dog, do you have pets too? *arXiv preprint arXiv:1801.07243*, 2018.
- [80] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*, 2019.
- [81] Yangjun Zhang, Pengjie Ren, and Maarten de Rijke. Improving background based conversation with context-aware knowledge pre-selection. *arXiv preprint arXiv:1906.06685*, 2019.
- [82] Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W Bruce Croft. Towards conversational search and recommendation: System ask, user respond. In *Proceedings of the 27th acm international conference on information and knowledge management*, pages 177–186, 2018.

- [83] Yinhe Zheng, Guanyi Chen, Minlie Huang, Song Liu, and Xuan Zhu. Personalized dialogue generation with diversified traits. *arXiv preprint arXiv:1901.09672*, 2019.
- [84] Peixiang Zhong, Chen Zhang, Hao Wang, Yong Liu, and Chunyan Miao. Towards persona-based empathetic conversational models. *arXiv preprint arXiv:2004.12316*, 2020.