Proactive Dialogue Systems and Conversational AI: A Survey

www.surveyx.cn

Abstract

Proactive dialogue systems represent a significant evolution in human-computer interaction, transcending the limitations of traditional rule-based frameworks by anticipating user needs and guiding conversations toward specific objectives. This survey explores the architecture, components, and methodologies that underpin these systems, emphasizing their role in enhancing interaction efficiency and user satisfaction. Key advancements include the integration of lifelong learning capabilities, cognitive psychology principles, and deep learning techniques, which collectively enable more personalized and effective dialogues. Applications span customer service, healthcare, and education, highlighting the systems' versatility in various domains. Despite these advancements, challenges persist, particularly in developing context-independent systems and improving overall system performance. Future research directions focus on enhancing model adaptability, user interaction, and ethical considerations, aiming to further refine the capabilities of proactive dialogue systems. The survey concludes by underscoring the importance of integrating planning and intention in dialogue systems to facilitate genuine collaborative interactions, suggesting that these innovations promise to redefine the landscape of human-computer interaction by offering more personalized, efficient, and engaging experiences across diverse applications.

1 Introduction

1.1 Significance in Human-Computer Interaction

Proactive dialogue systems have revolutionized human-computer interaction by overcoming the constraints of traditional rule-based frameworks, which often depend on linear processes and expensive labeled datasets [1]. These systems not only initiate meaningful dialogues but also enhance interaction efficiency and user satisfaction by anticipating user needs and steering conversations toward specific goals [2]. The integration of lifelong and continual learning capabilities enables these systems to adapt and personalize interactions over time [3]. Furthermore, applying cognitive psychology principles empowers these systems to influence user opinions and behaviors, thereby facilitating more effective and persuasive interactions [4]. Proactive communication is vital in fostering cooperation between humans and robots, which is essential for effective human-computer interaction [5].

Evaluating conversational quality is crucial for assessing both user experience and system performance, which are pivotal for advancing human-computer interaction [6]. However, challenges persist in the development of dialogue systems that are generic and context-independent, which are necessary for enhancing user interaction across diverse applications [7]. Addressing these challenges is essential for improving the accuracy and effectiveness of proactive dialogue systems [8]. Enhancing overall system performance in pipeline goal-oriented dialogue systems is critical, as current methodologies often focus on individual component improvements rather than holistic system performance [9].

Proactive dialogue systems personalize user experiences by considering user-specific situations, thoughts, and emotions over extended periods [10]. This personalization is particularly beneficial in

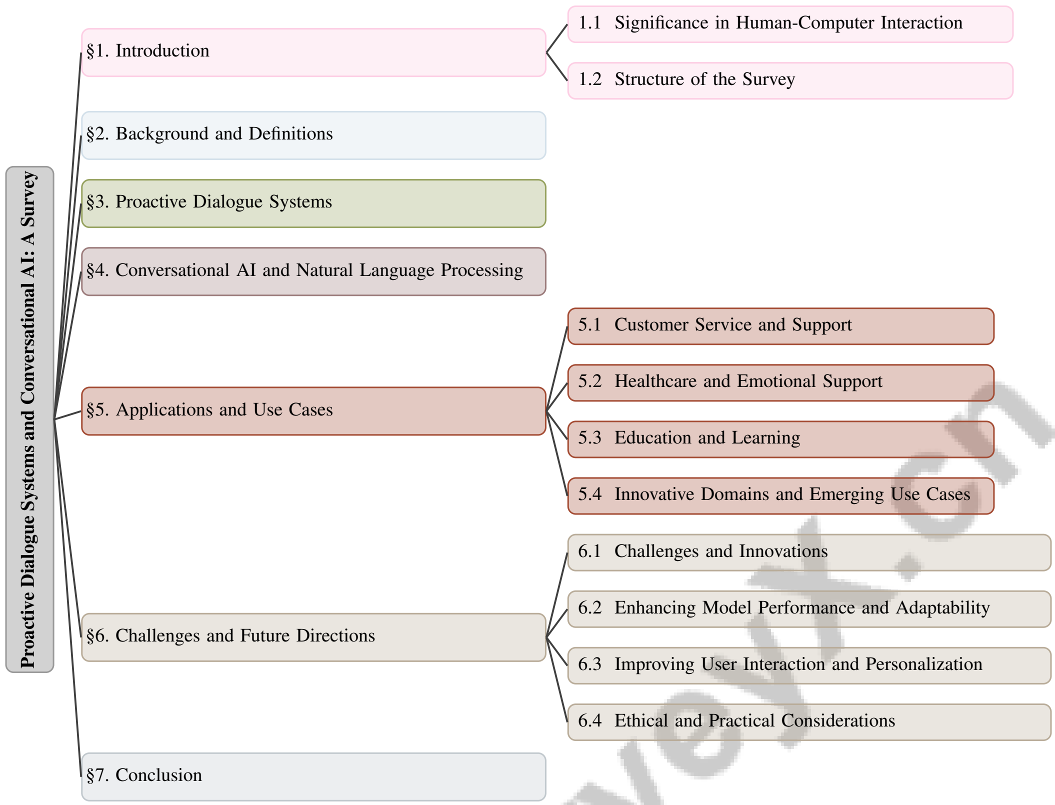


Figure 1: chapter structure

applications such as recommendation systems and customer service, where the creation of human-like conversational agents is a primary objective [11]. Effective dialogue management is essential for navigating breaches of conversational norms, highlighting its significance in enhancing human-robot interaction [12]. The evolution of chatbots in human-computer communication underscores their role in bridging knowledge gaps and improving performance in areas such as emotion perception, factuality, and informativeness. Despite the popularity of conversational agents, challenges in non-goal-oriented conversations remain [13]. There is an urgent need for standardized evaluation procedures for chat-oriented dialogue management systems to improve human-computer interaction [14]. Proactive dialogue systems enhance interactions by increasing the relevance and informativeness of responses [15].

The advancement of dialogue systems has been significantly propelled by deep learning techniques, which enhance their capabilities in managing complex multi-turn conversations. Nonetheless, the absence of planning in target-oriented conversational tasks presents a challenge, particularly in the context of goal-oriented AI agents [16]. The development of machine learning-driven spoken dialogue systems for goal-oriented interactions, especially for children, demonstrates the potential of these systems to enhance human-computer interaction [17]. User satisfaction estimation is another critical aspect, reflecting whether users' needs are met, typically inferred from dialogue acts [18]. The proactive guidance of conversations toward specific targets by goal-directed dialogue systems further emphasizes their importance in human-computer interaction [19]. Enhancing group goal-oriented discussions through conversational agents also underscores the necessity for frameworks that support such applications [20].

Communication breakdowns and loss of engagement in spoken dialogue systems highlight the need for breakthroughs to enhance user engagement [21]. Addressing Grice's maxims of conversation can improve natural language generation, particularly in generating referring expressions [22]. A comprehensive understanding of the concepts, issues, and technologies related to Conversational AI is essential for bridging knowledge gaps among researchers and graduate students in NLP and AI [23]. The lack of trust in conversational AI systems restricts their application in complex tasks requiring higher levels of cooperation [24]. Target-guided response generation is crucial for proactive control in

AI-generated, for reference only.

conversations, aiming to improve user experience and manage challenging situations [25]. Aligning large language model behavior with human logic enhances interpretability and interaction [26]. Automated dialogue systems often struggle to comprehend user emotions and provide empathetic feedback, which negatively impacts human-computer interaction [27]. Proactive dialogue systems strive to lead conversations toward goal-oriented topics while maintaining user satisfaction [28]. Teaching conversational agents to respond to problematic content in accordance with social norms can significantly improve user interactions and safety [29]. Effective communication between humans and intelligent agents in visual dialogue is vital for enhancing dialogue accuracy, particularly in the context of human answer mistakes [30].

The core challenge of generating coherent and contextually relevant dialogues in conversational agents is critical for achieving effective human-computer interaction [31]. The development of conversational robots like ERICA, capable of engaging in human-like dialogue, underscores the significance of attentive listening and job interview simulations in enhancing interactions [32]. The ability to learn efficient dialogue management from data with minimal manual intervention highlights the importance of proactive dialogue systems in improving human-computer interaction [33]. Target-oriented proactive dialogue systems enhance user engagement and coherence by directing conversations toward predetermined targets [34]. Goal-oriented dialogue systems assist users in achieving specific objectives within closed domains, thereby improving human-computer interaction [35]. The limitations of traditional task-oriented and search-oriented dialogue systems necessitate the development of proactive systems capable of engaging users in extended conversations [36]. Predicting user satisfaction in spoken dialogue systems, particularly in commercial systems like DuerOS, addresses challenges such as ASR errors and NLU misunderstandings that affect user experience [37]. Enhancing resource allocation in large-scale networks requires breakthroughs for improved scalability and efficiency [38]. The Duplex Conversation system exemplifies enhanced user experience through natural, human-like dialogue facilitated by effective turn-taking and responsive backchanneling [39]. Recent surveys highlight the significant role of deep learning-based dialogue systems in enhancing human-computer interaction [40]. The issue of proactivity in Information-Seeking Dialogue (ISD) agents, which have traditionally exhibited reactive behavior, is addressed in recent research, emphasizing the necessity for more proactive approaches in dialogue systems [41]. The psycholinguistic characteristics of dialogue systems are crucial for better interaction, as current systems often fail to resonate with users' psychological states and communication styles [42].

1.2 Structure of the Survey

This survey is systematically organized to provide a comprehensive examination of proactive dialogue systems and conversational AI. It begins with an introduction emphasizing their significance in enhancing human-computer interaction through anticipation of user needs and goal-setting. The subsequent section, Background and Definitions, offers an overview of core concepts and definitions relevant to proactive dialogue systems, ensuring clarity on terms such as 'proactive-dialogue', 'conversational AI', 'goal-oriented', and 'natural language processing'.

Following this, the survey delves into Proactive Dialogue Systems, exploring their architecture, components, and mechanisms for anticipating user needs and facilitating goal-oriented conversations. This section also examines knowledge integration and management within these systems. The next section, Conversational AI and Natural Language Processing, reviews advancements in dialogue generation models, AI technique integration, and the role of large language models in enhancing conversational AI.

Applications and Use Cases are then discussed, highlighting the impact of proactive dialogue systems across various domains such as customer service, healthcare, education, and emerging innovative areas. The survey thoroughly examines the Challenges and Future Directions in developing dialogue systems, addressing significant obstacles such as the limitations of current interaction paradigms, the need for improved model performance through goal-oriented prompt engineering, and the ethical implications of anthropomorphic assumptions in AI design. It proposes future research avenues, including enhancing user interaction through proactive response mechanisms, exploring multi-modal and cross-cultural negotiation scenarios, and addressing the complexities of long-term context in dialogue management, ultimately aiming to foster more effective and ethically responsible dialogue systems [43, 26, 44, 45].

Finally, the Conclusion summarizes the key points discussed throughout the paper and suggests areas for future research and development. This structured approach ensures a thorough understanding of the current landscape and future prospects of proactive dialogue systems and conversational AI, as categorized in the comprehensive overview provided by [46].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Definitions of Core Concepts

Proactive dialogue systems are designed to engage users by anticipating their needs and guiding conversations towards specific objectives, known as 'proactive-dialogue.' This approach is particularly effective in handling interactions with non-cooperative behaviors, preventing conversations from deviating [28]. A key element is 'target-guided response generation,' which utilizes commonsense knowledge graphs to enhance interaction relevance and coherence [25].

Conversational AI refers to technologies enabling machines to interact with humans using natural language, encompassing both task-oriented and open-domain dialogue systems. These systems employ natural language processing (NLP) to maintain engagement and coherence across various contexts. The integration of 'emotionally intelligent dialogue generation models' (EIDGM) improves conversational AI by detecting user emotions and generating empathetic responses, thus increasing user satisfaction [27].

'Goal-oriented dialogue systems' assist users in completing specific tasks through structured, multi-turn conversations, crucial for discerning user intent and maintaining coherence. They often involve backend service communication to fulfill tasks, with natural language understanding (NLU) tasks like slot filling (SF) and intent classification (IC) being essential. 'Goal-oriented prompting' enhances large language model (LLM) performance across tasks, ensuring focused and effective interactions [26].

In dialogue safety, 'prosocial behavior' involves actions that benefit others, essential for responding to potentially unsafe user inputs [29]. Additionally, intelligent agents must identify and correct human interlocutor errors, particularly in visual dialogue interactions, to maintain conversation integrity and accuracy [30]. Understanding these core concepts is vital for advancing proactive dialogue systems and conversational AI, forming a foundation for systems that effectively engage users and achieve conversational goals [41]. This understanding supports the evolution of deep conversational recommender systems (DCRS), which combine user intention comprehension with personalized recommendations in a conversational framework [47].

2.2 Evolution of Dialogue Systems

The evolution of dialogue systems in AI has undergone significant changes marked by technological advancements. Initially, rule-based systems dominated, relying on handcrafted rules and manual configurations for interaction management, laying the groundwork for future human-computer communication developments [48]. As AI advanced, statistical models introduced probabilistic approaches, enhancing dialogue management flexibility and performance [23].

The differentiation between task-oriented and open-domain dialogue systems marked a critical phase in this evolution. Task-oriented systems focused on specific goals, using techniques like slot filling and intent classification for structured interactions [49]. In contrast, open-domain systems engaged users in broader conversations, often struggling with coherence and relevance across diverse topics [31].

The integration of pre-trained language models into dialogue systems was a pivotal shift, enhancing natural language understanding and generation capabilities. This integration improved both task-oriented and open-domain systems, leveraging large language models for better conversational fluency and contextual awareness [50]. Despite these advancements, dialogue system evaluations often remained component-focused, highlighting the need for holistic assessment methods considering overall system performance [51].

Contemporary dialogue systems have adopted deep learning techniques, significantly improving their ability to manage complex, multi-turn interactions. This evolution includes end-to-end neural network approaches, offering scalability and adaptability beyond traditional methods [40]. However, challenges persist in multimodal systems, particularly in goal-oriented tasks like Visual Dialogue, which current generative language models inadequately address [52].

The ongoing evolution of dialogue systems underscores the need for comprehensive frameworks integrating diverse methodologies, from traditional handcrafted techniques to modern data-driven approaches [53]. As dialogue systems progress, combining these varied approaches promises to enhance their effectiveness in engaging users and achieving conversational goals.

In recent years, the advancement of proactive dialogue systems has garnered significant attention within the field of human-computer interaction. These systems leverage sophisticated methods to anticipate user needs and enhance overall engagement. Figure 2 illustrates the hierarchical structure of proactive dialogue systems, detailing their architecture and components. The figure categorizes these systems into key paradigms and features, while also presenting innovative frameworks that emphasize predictive methodologies and user interaction classification. Furthermore, it highlights the critical role of dialogue management and adaptive strategies in achieving conversational objectives, thereby underscoring their importance in fostering meaningful user experiences. This comprehensive representation serves as a foundation for understanding the complexities involved in the design and implementation of effective proactive dialogue systems.

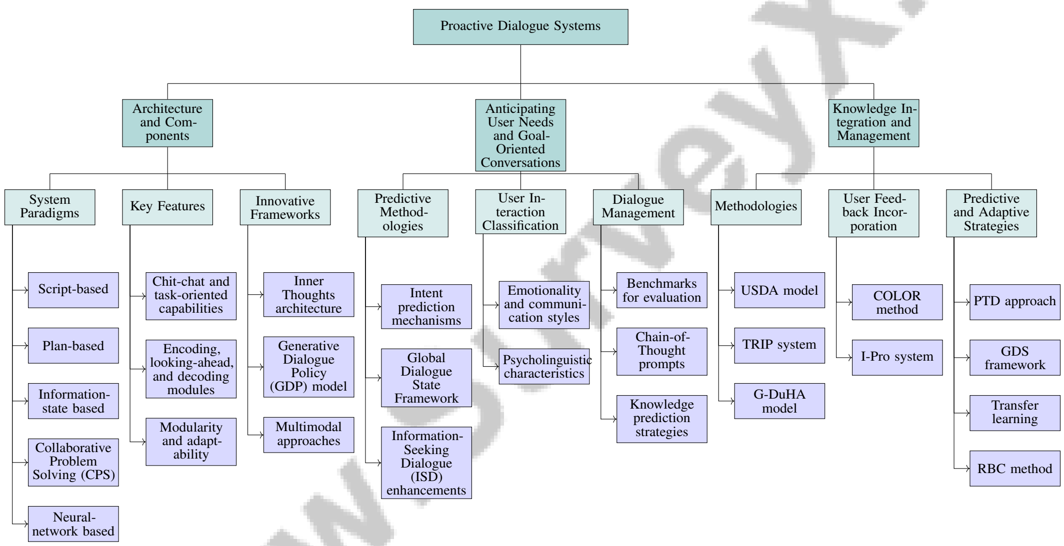


Figure 2: This figure illustrates the hierarchical structure of proactive dialogue systems, detailing their architecture and components, methods for anticipating user needs, and strategies for knowledge integration and management. The systems are categorized into key paradigms, features, and innovative frameworks, with emphasis on predictive methodologies and user interaction classification. The diagram also highlights the role of dialogue management and adaptive strategies in enhancing user engagement and achieving conversational objectives.

3 Proactive Dialogue Systems

3.1 Architecture and Components of Proactive Dialogue Systems

Proactive dialogue systems are meticulously architected to enable dynamic, goal-driven interactions through an amalgamation of advanced components and methodologies. These systems are classified into paradigms such as script-based, plan-based, information-state based, collaborative problem solving (CPS), and neural-network based approaches [53]. A key feature is the integration of both chit-chat and task-oriented capabilities, exemplified by the Unified Dialogue System (UniDS), which fuses these dialogue types via an auto-regressive language model [54]. Similarly, the Neural Network-based Goal-oriented Dialogue System (NN-GDS) utilizes sequence-to-sequence learning, belief tracking, and database querying to manage complex interactions effectively [55].

These systems often incorporate encoding, looking-ahead, and decoding modules, as demonstrated by the end-to-end dialogue model (E2E-DM), which anticipates user needs to enhance proactive engagement [33]. Systems like ERICA, an autonomous android, employ advanced dialogue management and response generation techniques to simulate human-like interactions, showcasing sophisticated dialogue strategies [32].

Innovative frameworks such as the Inner Thoughts architecture enable AI to autonomously predict user needs through stages like trigger, retrieval, thought formation, evaluation, and participation [56]. The Generative Dialogue Policy (GDP) model uses recurrent neural networks (RNNs) to develop adaptable dialogue policies, highlighting the necessity of flexible strategies [57]. Multimodal approaches, such as the Duplex Conversation method, enhance interaction management by combining audio and text inputs to improve turn-taking and manage user interruptions [39]. The Discourse Relation Dialogue Model (DRDM) utilizes discourse relations and structured ontology to manage diverse conversational contexts [36].

These architectures exhibit modularity and adaptability, integrating advanced techniques like knowledge-augmented dialogue systems that combine knowledge retrieval with language models to improve action prediction and dialogue management [58]. Frameworks such as Diplomat exemplify this modularity by streamlining the development of conversational agents for group discussions [20].

As illustrated in Figure 3, the hierarchical structure of proactive dialogue systems categorizes them into dialogue paradigms, models, and innovative frameworks, each contributing uniquely to enhancing conversational capabilities.

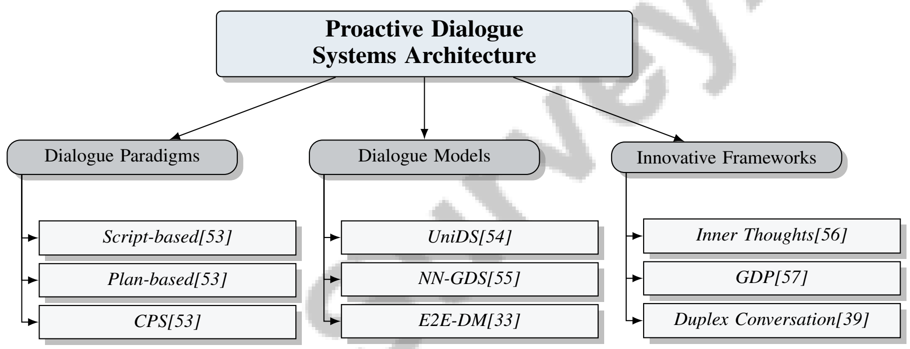


Figure 3: This figure illustrates the hierarchical structure of proactive dialogue systems, categorizing them into dialogue paradigms, models, and innovative frameworks, each contributing uniquely to enhancing conversational capabilities.

3.2 Anticipating User Needs and Goal-Oriented Conversations

Proactive dialogue systems excel in anticipating user needs and fostering goal-oriented conversations through sophisticated predictive methodologies and dynamic frameworks. These systems surpass traditional reactive models by utilizing advanced intent prediction mechanisms, enabling seamless navigation toward specific objectives. A significant challenge is maintaining response coherence and managing dialogue context, essential for anticipating user needs [40].

The Global Dialogue State Framework exemplifies a methodology allowing dialogue agents to assess the likelihood of achieving desired outcomes and adjust responses accordingly, enhancing adaptability to evolving conversational contexts [59]. Reconceptualizing proactivity in Information-Seeking Dialogue (ISD) involves adding Follow-up Questions (FQ) or Additional Information (AI) related to the user's initial query, enriching the dialogue and anticipating further needs [41].

Accurate classification of user interactions based on emotionality and communication styles is crucial for enhancing engagement [42]. Understanding the psycholinguistic characteristics of user interactions allows systems to tailor responses to align with user expectations and emotional states, thereby improving satisfaction.

The development of benchmarks highlights the need for innovative dialogue management approaches, as traditional decoding strategies are inadequate for multimodal dialogue systems [52]. These benchmarks provide a structured environment for evaluating systems' strengths and weaknesses, ensuring effective anticipation of user needs and facilitation of targeted conversations. Advanced techniques such as Chain-of-Thought prompts and knowledge prediction strategies enable these systems to achieve interaction objectives more effectively, significantly improving dialogue quality and relevance [60, 61, 62, 41].

3.3 Knowledge Integration and Management

Knowledge integration and management are critical in proactive dialogue systems, ensuring interactions remain contextually relevant and coherent. These systems employ various methodologies to manage and integrate knowledge effectively, optimizing user engagement and achieving conversational objectives. The USDA model highlights the importance of modeling relationships between dialogue content and acts, enhancing user satisfaction predictions through sequential transitions [18].

The G-DuHA model exemplifies goal-oriented dialogue maintenance by accurately representing interlocutors' roles, thereby improving dialogue quality and relevance [63]. This is supported by systems like TRIP, which utilize dual Transformer decoders to ensure coherence and relevance in dialogue generation, integrating knowledge from user profiles and domain knowledge graphs [34].

Incorporating user feedback is essential, as shown by the COLOR method, which enhances dialogue planning and anticipates user needs through proactive strategies [19]. The I-Pro system innovatively derives a learned goal weight from various factors, enabling strategy adaptation based on conversational context and user behavior [28].

The PTD approach leverages predictions about user and agent behavior to make contextually aware decisions [64]. This predictive capability is complemented by the GDS framework, integrating both local and global dialogue state tracking to enhance dialogue progression awareness [59]. Transfer learning is a critical technique, improving policy learning in Goal-Oriented Dialogue Systems by leveraging domain similarities for enhanced knowledge integration [35]. The RBC method underscores dynamic knowledge management by adapting to unrecognized utterances, facilitating effective clustering and analysis [65].

These methodologies underscore the necessity of robust knowledge integration and management in proactive dialogue systems, enabling dynamic, contextually relevant interactions that enhance user experience and achieve specific conversational goals. Incorporating proactive elements such as Follow-up Questions or Additional Information in response generation further enriches dialogues by anticipating and addressing future user needs [41].

4 Conversational AI and Natural Language Processing

Category	Feature	Method
Advancements in Dialogue Generation Models	Representation and Management	DC[39], QDLM[66]
	Predictive and Transfer Techniques	PTD[64], TL-GO[35]
	Versatile Dialogue Systems	UniDS[54]
	Goal-Oriented Dialogue	TRIP[34], GDS[59]
Integration of AI Techniques in Dialogue Systems	Adaptive Dialogue Management	FC[67], RBC[65]
	Emotional and Empathetic Interaction	EIDGM[27]
	Error Handling and Recognition	PHMT[30]
Role of Large Language Models (LLMs)	Real-Time Capabilities	DUO[68]
	Interaction Management	DRDM[36], MCM[69]

Table 1: This table provides a comprehensive overview of recent advancements in dialogue generation models, the integration of AI techniques in dialogue systems, and the role of large language models. It categorizes various features and methods that have significantly enhanced the capability, adaptability, and user interaction quality of conversational AI systems.

The convergence of computational advancements and natural language processing has catalyzed significant progress in conversational AI, particularly enhancing dialogue systems. Table 1 presents a detailed classification of the latest developments in dialogue systems, highlighting key methodologies and innovations that drive improvements in conversational AI. This section explores key developments in dialogue generation models that elevate interaction quality and user engagement, thereby optimizing the efficacy and adaptability of conversational agents across diverse contexts.

4.1 Advancements in Dialogue Generation Models

Recent advancements in dialogue generation models have significantly improved conversational systems, enabling the production of coherent, contextually appropriate, and engaging interactions. Deep learning techniques have been pivotal, allowing models to comprehend complex user inputs and generate nuanced responses, thus enriching user experience [40]. Noteworthy advancements include the TRIP method, which surpasses baseline models in generating target-oriented dialogues, demonstrating the efficacy of advanced planning techniques [34]. The PTD framework distinguishes itself by using two prediction models to simulate future dialogue states, facilitating more informed decision-making compared to traditional methods [64].

Reinforcement learning integration marks another critical advancement, with Transfer Learning enhancing dialogue management adaptability to new domains and tasks [35]. The QDLM model employs quantized dialogue representations to boost performance, highlighting the importance of innovative data representation in dialogue generation [66]. Unified dialogue systems like UniDS exemplify the trend towards versatile models that generate responses for both chit-chat and task-oriented contexts [54]. The Duplex Conversation system has improved response latency and user interaction, outperforming baseline methods across multiple metrics, illustrating the impact of optimized interaction management [39].

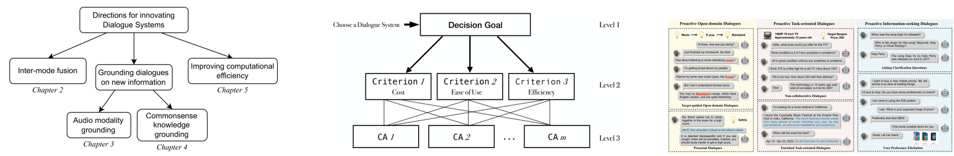
The GDS framework introduces a task-specific progression function that enhances dialogue generation by providing turn-level estimates of task success, offering precise control over interaction flow [59]. Additionally, psycholinguistic benchmarks focusing on contextually relevant traits enrich dialogue systems by enabling more personalized and adaptable interactions [42]. These advancements, spanning deep learning, reinforcement learning, and sophisticated planning techniques, have markedly improved the adaptability, coherence, and effectiveness of conversational systems. The Confirm-it strategy has emerged as a leading approach regarding accuracy and hallucination rates, highlighting trade-offs between lexical variety and precision [52].

4.2 Integration of AI Techniques in Dialogue Systems

The integration of artificial intelligence techniques within dialogue systems has significantly enhanced their capabilities, enabling nuanced, context-aware, and adaptive interactions. A key development is the incorporation of emotional intelligence technology, which allows systems to generate empathetic responses through advanced natural language processing, thereby improving user satisfaction and engagement [27]. Prosocial dialogue frameworks exemplify this integration, where datasets are created from human-AI collaborative interactions. Models like GPT-3 generate potentially unsafe utterances, while human crowdworkers provide prosocial responses, training the system to handle sensitive topics appropriately [29].

Furthermore, recognizing and addressing common human errors in interactions enhances dialogue system reliability [30]. AI-driven clustering algorithms, such as those employed by the RBC algorithm, play a vital role in improving task-oriented dialogue systems by effectively categorizing unrecognized user inputs, leading to more accurate dialogue management [65]. Collectively, these AI techniques enhance dialogue system adaptability and effectiveness, paving the way for future innovations in conversational AI. By integrating emotional intelligence, prosocial frameworks, error recognition, and clustering algorithms, dialogue systems can engage users in more meaningful, informative, and contextually relevant interactions across various domains. This integration fosters the development of empathetic dialogue systems capable of recognizing and responding to user emotions while proactively introducing relevant information, improving the overall quality and effectiveness of user interactions [8, 27, 62].

As illustrated in Figure 4, the integration of AI techniques in dialogue systems is a rapidly evolving field that leverages advancements in conversational AI and natural language processing to enhance human-computer interactions. This exploration focuses on three key aspects: innovation directions, decision-making criteria, and dialogue typologies. The first aspect, "Directions for Innovating Dialogue Systems," is represented by a flowchart delineating essential pathways for advancement, emphasizing inter-mode fusion, grounding dialogues on novel information, and enhancing computational efficiency. The second aspect, "Decision Goal and Criteria for Choosing a Dialogue System," outlines a structured decision-making process, highlighting critical evaluation criteria such as cost, ease of use, and efficiency for selecting the most suitable dialogue system. Lastly, the "Dialogue Types and Their



(a) Directions for innovating Dialogue Systems[67]

(b) Decision Goal and Criteria for Choosing a Dialogue System[6]

(c) Dialogue Types and Their Characteristics[46]

Figure 4: Examples of Integration of AI Techniques in Dialogue Systems

Characteristics" aspect provides a comparative analysis of various dialogue types, categorizing them into proactive open-domain, task-oriented, and information-seeking dialogues, each with distinct characteristics and subcategories. Together, these elements illustrate a comprehensive approach to integrating AI techniques in dialogue systems, aiming to create more sophisticated, efficient, and user-friendly conversational agents [67, 6, 46].

4.3 Role of Large Language Models (LLMs)

Large Language Models (LLMs) have become transformative components in conversational AI, significantly enhancing the naturalness, coherence, and strategic capabilities of dialogue systems. These advancements have reshaped dialogue systems, enabling the management of complex interactions and delivering human-like conversational experiences [50]. Models like OpenAI's GPT-4 have demonstrated superior performance in tasks such as proactivity prediction, showcasing their potential in anticipating user needs and effectively guiding conversations [70].

The introduction of the Cskills benchmark has further enriched the anthropomorphic qualities of LLMs, enhancing their engagement in dialogues by mimicking human communication styles [71]. Despite these advancements, challenges persist, particularly in strategic decision-making during non-collaborative scenarios. Models like ProCoT have shown improved abilities in asking clarification questions and managing topic transitions but continue to struggle with complex strategic decisions [72].

Innovative approaches such as the DUplex decOding (DUO) method facilitate real-time processing and output generation, enhancing the fluidity and responsiveness of conversational interactions [68]. The integration of structured control mechanisms over LLM outputs has also improved dialogue stability and naturalness, providing a more controlled and consistent user experience [69]. LLMs also play a crucial role in optimizing interaction strategies and providing contextual expertise through the integration of large ontologies and discourse relations, facilitating more natural and contextually relevant interactions [36]. However, a notable performance gap between LLMs and human users remains, as evidenced by recent benchmarks involving state-of-the-art models like ChatGPT and LLaMA [73].

5 Applications and Use Cases

Proactive dialogue systems have advanced beyond traditional roles, revolutionizing user interactions and service delivery across diverse sectors, including customer service, healthcare, education, and emerging domains.

5.1 Customer Service and Support

In customer service, proactive dialogue systems enhance interaction efficiency and user satisfaction through advanced task-oriented frameworks, predicting user needs and guiding conversations toward objectives [51]. The SlugBot system exemplifies improved engagement in open-domain dialogues, demonstrating utility in customer service [36]. Systems utilizing datasets from real human dialogues, such as the Expert Live Chat, ensure realistic interactions [74]. Personalization based on user profiles is crucial, as seen in benchmarks for end-to-end trained systems in goal-oriented conversations

[75]. Transitioning from casual to task-oriented dialogue is vital for training sales agents, ensuring goal-driven interactions [76].

Incorporating non-verbal cues provides a human-like experience [77], while systems in online shopping assist with product inquiries and recommendations [78]. Elements of politeness and positivity further enhance satisfaction [79]. The Policy2Target framework optimizes strategies in FinTech for debt collection, exemplifying strategic improvement [80]. Prosocial frameworks address harmful comments positively [29], and the E2E-IQRS framework enhances agent efficiency with proactive questions [44]. These applications set new standards in customer service, impacting domains like human resources and career coaching [81].

5.2 Healthcare and Emotional Support

In healthcare, proactive dialogue systems enhance patient interactions and emotional support. They create realistic clinical dialogues for diagnostic information collection [82] and employ emotionally intelligent systems to recognize patient emotions, fostering supportive environments [27]. These systems are vital in mental health, leveraging datasets like the 800 two-session Longitudinal Dialogues for emotional support [10], and engage young users in education by addressing unique interaction patterns [17]. Surveys highlight their impact on user experience and emotional support [83], with systems like ERICA offering insights into effective communication in social and healthcare settings [32].

5.3 Education and Learning

Proactive dialogue systems enhance education by providing personalized, interactive support. Utilizing advanced NLP and machine learning, they create engaging environments tailored to individual learning styles [27]. These systems guide students through complex subjects, offering real-time feedback and identifying areas of struggle for targeted assistance. Their role in educational games and simulations exemplifies interactive learning [17]. They facilitate collaborative learning by promoting inclusivity and simulating real-world scenarios, bridging theory and practice [32]. Emotionally intelligent systems enhance engagement by recognizing students' emotions and providing encouragement [27].

5.4 Innovative Domains and Emerging Use Cases

Proactive dialogue systems are increasingly applied in innovative domains, showcasing their expanding capabilities. Generative models enhance personalization and adaptability in conversational agents [84]. Seamless domain transitions in multi-domain settings offer integrated experiences [85], with datasets emphasizing personalized interactions and proactive management [61]. The evolution of personal assistants to social bots marks a significant innovation, broadening applications in social domains [86]. Low-resource multilingual systems highlight potential across diverse contexts [87]. The Alexa Prize competition advances open-domain conversational AI [13], while conversational programming systems suggest innovative applications in software development [88].

Future applications of systems like GOMA in real-world settings highlight potential in proactive dialogue for cooperative tasks [5]. The DUO method's multimodal extension enhances applicability beyond text [68], and annotated dialogue data benchmarks ensure scalable training across domains [89]. Future research will focus on deploying systems like I-Pro in real-world settings to refine effectiveness and adaptability [28]. These emerging use cases demonstrate the versatility of proactive dialogue systems in complex scenarios, promising to redefine human-computer interaction with personalized, efficient, and engaging experiences.

6 Challenges and Future Directions

6.1 Challenges and Innovations

Proactive dialogue systems face significant challenges due to the complexities of human dialogue and model limitations. A critical issue is managing the vast response possibilities, which can degrade system performance. The Quantized Dialogue Language Model (QDLM) addresses this by efficiently handling the extensive response space [66]. The scarcity of labeled dialogue data further limits adaptability and generalization across domains [35]. Unrecognized utterances pose another challenge, often causing system errors; the RBC algorithm enhances processing unexpected inputs [65]. Developing a unified model capable of handling various tasks without performance loss is difficult, especially in contexts requiring specialized dialogue skills [54]. Additionally, reliance on curated ontologies can lead to conversational gaps [36].

Current models often struggle with generalizing across complex, real-world dialogues due to inadequate datasets and costly labeling processes, limiting commercial applicability [37]. High-quality audio and text input dependence complicates real-world applications, as input quality declines can significantly affect performance [39]. Challenges also include ensuring response coherence, managing dialogue context, and integrating external knowledge [40].

Innovations offer promising solutions. Improved adaptability to changing conditions and resource utilization enhances overall network performance [38]. Progression-aware autonomous dialogue agents, initially limited to single tasks and datasets, show potential for broader applications [59]. Knowledge augmentation and multitasking strategies improve systems' abilities to manage inter-mode dependencies between task-oriented and open-domain dialogues. Modular frameworks enhance transparency and performance, addressing model obfuscation from joint modeling and integrating diverse methodologies [81]. Token-level generation methods aim to generate semantically coherent responses while reducing computation times.

A notable limitation is the dependence on underlying language models, which may introduce biases or inconsistencies in response quality [41]. The reliance on clear inter-annotator agreement in benchmarks may restrict applicability to more ambiguous real-world interactions [42]. Current decoding strategies for multimodal interactions require further exploration to fully capture these dynamics [52].

6.2 Enhancing Model Performance and Adaptability

Enhancing dialogue models' performance and adaptability is crucial for effective user engagement across varied contexts and tasks. Converting qualitative success criteria into quantifiable metrics, such as turn-level dialogue evaluation methods, facilitates real-time performance assessments, providing a framework for refining system interactions [90]. The QDLM improves model performance by narrowing the prediction space, focusing on relevant utterance clusters to enhance response accuracy and coherence [66].

Future research should integrate neural and plan-based approaches to develop robust evaluation frameworks and leverage large language models for data generation and annotation [53]. Incorporating Dialogue Flow Control Prompt (DFCP) and Turn-Take Control Prompt (TTCP) into a singular prompt may enhance interaction fluidity and naturalness [69].

Data augmentation techniques, such as paraphrase generation and entity extraction, are vital for improving model adaptability by increasing training data diversity and volume. This approach broadens scenarios models can handle, enhancing robustness to user input variations [91]. Future research could focus on addressing rare concepts and enhancing response diversity, tackling current dialogue systems' limitations [92].

Scaling models for larger domains and improving resilience to noisy inputs remain critical challenges in performance and adaptability. Addressing these issues is essential for advancing dialogue systems capable of functioning effectively in real-world environments [55]. Refining low engagement detection and integrating findings into improved autonomous systems will enhance user interaction, particularly in multi-user scenarios [21].

6.3 Improving User Interaction and Personalization

Enhancing user interaction and personalization in dialogue systems is crucial for creating engaging and responsive experiences. Modular designs, like the Diplomat conversational agent framework, offer flexibility and adaptability to diverse user needs and contexts [20]. The Intent Space Model further enhances adaptability by incorporating new user intents without complete retraining, dynamically improving personalization [93].

Method Name	Modular Design	Dynamic Adaptability	Multimodal Integration
N/A[20]	Modular Agent Features	Update Parameters Dynamically	-
ISM[93]	Intent Space Model	Adding Unseen Intents	-
GDP[57]	-	-	-
TBM[37]	Flexible Frameworks	New User Intents	Structured And Text
DC[39]	Flexible Turn-taking	Dynamic Parameter Updates	Multimodal Inputs
GDS[59]	-	-	-
PDG[41]	-	-	-
TL-GO[35]	-	-	-
TRIP[34]	Flexible Frameworks	New User Intents	Various Data Types
PTD[64]	Flexible Frameworks	Update Parameters Dynamically	Various Data Types
GOIS[17]	Rasa Framework	Dialog Adaptation Techniques	Multimodal Conversational System
DRDM[36]	Structured Ontology	Mixed Initiative Interactions	Datasets From Reddit

Table 2: This table presents a comparative analysis of various dialogue system methods with respect to their modular design, dynamic adaptability, and multimodal integration capabilities. It highlights the unique features and capabilities of each method, providing insights into their potential for enhancing user interaction and personalization in dialogue systems.

Table 2 provides a comprehensive comparison of different dialogue system methods, focusing on their modular design, dynamic adaptability, and multimodal integration, which are crucial for improving user interaction and personalization. The GDP model enhances user interaction and personalization by generating multiple dialogue acts with parameters simultaneously in task-oriented dialogue systems [57]. Research should focus on improving context incorporation, response personalization, and addressing generic response generation [31]. Refining benchmarks to include diverse scenarios and improving model robustness in real-world applications is also crucial [94].

The TBM approach utilizes structured and unstructured data to improve user satisfaction predictions, adapting to real-world dialogue complexities [37]. Future work may explore reinforcement learning for dynamic parameter updates to enhance interaction and adaptability [39]. Investigating user feedback's role in improving dialogue systems is also essential [40].

The GDS framework offers a promising avenue for enhancing user interaction and personalization by managing dialogue progression across various tasks and datasets [59]. This approach improves user engagement and sustains conversations, leading to greater satisfaction in information-seeking dialogues [41]. Tailoring chatbot responses could lead to higher user satisfaction [42].

Strategies for improving user interaction include Transfer Learning to enhance Goal-Oriented chatbots' performance, especially in data-constrained environments [35]. Future research could enhance planning process robustness and model generalization across domains [34]. Integrating transformer-based models could further enhance prediction accuracy and decision-making [64].

Incorporating multimodal cues is crucial for maintaining engagement, particularly in interactions with children, where adaptation to unpredictable dialogue patterns is essential [17]. Improving response coherence and relevance through structured planning in dialogue generation is another effective strategy. Future research should enhance systems' ability to adapt to user emotions and investigate reinforcement learning techniques to refine dialogue strategies [36]. Further refinement of decoding strategies is necessary to better address multimodal dialogue challenges, balancing accuracy and linguistic richness [52].

6.4 Ethical and Practical Considerations

Deploying proactive dialogue systems and conversational AI introduces ethical and practical challenges that must be addressed for responsible and effective implementation. Protecting user data and privacy during processing is crucial for maintaining trust and transparency in AI interactions [65]. Safeguarding user data and respecting privacy is paramount as these systems become increasingly integrated into daily life.

Ethical considerations include protecting and compensating annotators who may encounter problematic content during data annotation. Measures should shield annotators from potential psychological harm and ensure fair compensation [65]. Additionally, deploying persuasive dialogue systems requires careful management to prevent bias and malicious discourse, ensuring beneficial purposes [59].

From a practical perspective, input data quality significantly influences dialogue systems' performance. Poor data quality can lead to suboptimal performance and impede effective resource allocation [38]. High-quality data inputs are necessary for optimal functionality and relevant interactions.

Developing standardized evaluation frameworks is essential for comprehensively assessing dialogue system performance. Current methods often fall short, relying on fragmented protocols lacking standardization. A synthesis of 20 studies highlights the limitations of automated and human evaluation techniques, emphasizing the need for a unified and rigorous framework. Establishing robust evaluation metrics will be crucial for advancing the field and ensuring systems meet desired performance and user satisfaction standards [45, 14, 95, 96, 97].

7 Conclusion

Proactive dialogue systems and conversational AI have significantly advanced human-computer interaction by enhancing dialogue naturalness, coherence, and goal-orientation. The incorporation of deep learning techniques and large language models has enabled the creation of fluent and contextually appropriate responses, particularly in goal-directed tasks. Transfer learning has further enhanced system performance in data-scarce settings, indicating promising future research paths in refining these processes. Innovative models like the Generative Dialogue Policy (GDP) have demonstrated superior performance over traditional methods, highlighting the potential for advanced dialogue management strategies. The UniDS system exemplifies the successful integration of task-oriented and open-domain dialogue capabilities, suggesting new research directions in dialogue transitions and system unification.

The role of emotional intelligence in dialogue systems is crucial for improving user interaction and satisfaction, indicating a need for deeper exploration in this area. Adaptive Resource Allocation Algorithms (ARAA) have proven effective in enhancing system efficiency and adaptability, affirming their relevance in practical applications. Future research should focus on expanding datasets and integrating multimodal information to strengthen proactive dialogue systems. Applying cognitive psychology principles may increase persuasiveness and broaden application scenarios. A comprehensive evaluation framework that considers both conversational quality and task completion is essential for optimizing user experience in task-oriented systems.

Effective dialogue systems must incorporate planning and intention to facilitate genuine collaborative interactions. Future work will aim to integrate these elements into fully end-to-end architectures. Enhancing capabilities such as model-based dialogue policies and supporting open-domain conversations will be critical. Proposed methodologies demonstrate significant improvements in performance and interpretability by integrating diverse dialogue skills into a cohesive system. Ongoing research should aim to reduce user reliance on corrections, potentially through automated feedback processes or more efficient learning algorithms, thereby improving system efficiency and user satisfaction. Refining evaluation methods and exploring social language variations could further enhance dialogue system efficacy. As these systems continue to evolve, they promise to transform human-computer interaction, offering more personalized, efficient, and engaging experiences across a wide range of applications.

References

[1] Jianfeng Liu, Feiyang Pan, and Ling Luo. Gochat: Goal-oriented chatbots with hierarchical reinforcement learning, 2020.

[2] Tian Xie, Xinyi Yang, Angela S. Lin, Feihong Wu, Kazuma Hashimoto, Jin Qu, Young Mo Kang, Wenpeng Yin, Huan Wang, Semih Yavuz, Gang Wu, Michael Jones, Richard Socher, Yingbo Zhou, Wenhao Liu, and Caiming Xiong. Converse: A tree-based modular task-oriented dialogue system, 2022.

[3] Sahisnu Mazumder and Bing Liu. Lifelong and continual learning dialogue systems, 2024.

[4] Mengqi Chen, Bin Guo, Hao Wang, Haoyu Li, Qian Zhao, Jingqi Liu, Yasan Ding, Yan Pan, and Zhiwen Yu. The future of cognitive strategy-enhanced persuasive dialogue agents: New perspectives and trends, 2024.

[5] Lance Ying, Kunal Jha, Shivam Aarya, Joshua B. Tenenbaum, Antonio Torralba, and Tianmin Shu. Goma: Proactive embodied cooperative communication via goal-oriented mental alignment, 2024.

[6] Ryan Fellows, Hisham Ihshaish, Steve Battle, Ciaran Haines, Peter Mayhew, and J. Ignacio Deza. Task-oriented dialogue systems: performance vs. quality-optima, a review, 2021.

[7] Maali Mnasri. Recent advances in conversational nlp : Towards the standardization of chatbot building, 2019.

[8] Yukun Ma, Khanh Linh Nguyen, Frank Z Xing, and Erik Cambria. A survey on empathetic dialogue systems. *Information Fusion*, 64:50–70, 2020.

[9] Zichuan Lin, Jing Huang, Bowen Zhou, Xiaodong He, and Tengyu Ma. Joint system-wise optimization for pipeline goal-oriented dialog system, 2021.

[10] Seyed Mahed Mousavi, Simone Caldarella, and Giuseppe Riccardi. Response generation in longitudinal dialogues: Which knowledge representation helps?, 2023.

[11] Zelin Dai, Weitang Liu, and Guanhua Zhan. Multiple generative models ensemble for knowledge-driven proactive human-computer dialogue agent, 2020.

[12] Maitreyee Tewari, Thomas Hellström, and Suna Bensch. Conversational norms for human-robot dialogues, 2021.

[13] Anu Venkatesh, Chandra Khatri, Ashwin Ram, Fenfei Guo, Raefer Gabriel, Ashish Nagar, Rohit Prasad, Ming Cheng, Behnam Hedayatnia, Angeliki Metallinou, Rahul Goel, Shaohua Yang, and Anirudh Raju. On evaluating and comparing open domain dialog systems, 2018.

[14] Sarah E. Finch and Jinho D. Choi. Towards unified dialogue system evaluation: A comprehensive analysis of current evaluation protocols, 2020.

[15] Behnam Hedayatnia, Karthik Gopalakrishnan, Seokhwan Kim, Yang Liu, Mihail Eric, and Dilek Hakkani-Tur. Policy-driven neural response generation for knowledge-grounded dialogue systems, 2020.

[16] Yosuke Kishinami, Reina Akama, Shiki Sato, Ryoko Tokuhisa, Jun Suzuki, and Kentaro Inui. Target-guided open-domain conversation planning, 2022.

[17] Saurav Sahay, Shachi H Kumar, Eda Okur, Haroon Syed, and Lama Nachman. Modeling intent, dialog policies and response adaptation for goal-oriented interactions, 2019.

[18] Yang Deng, Wenxuan Zhang, Wai Lam, Hong Cheng, and Helen Meng. User satisfaction estimation with sequential dialogue act modeling in goal-oriented conversational systems, 2022.

[19] Jian Wang, Dongding Lin, and Wenjie Li. Dialogue planning via brownian bridge stochastic process for goal-directed proactive dialogue, 2023.

[20] Kevin Hogan, Annabelle Baer, and James Purtilo. Diplomat: A conversational agent framework for goal-oriented group discussion, 2021.

[21] Haruki Kawai, Divesh Lala, Koji Inoue, Keiko Ochi, and Tatsuya Kawahara. Evaluation of a semi-autonomous attentive listening system with takeover prompting, 2024.

[22] Robert Dale and Ehud Reiter. The role of the gricean maxims in the generation of referring expressions, 1996.

[23] Michael McTear. *Conversational ai: Dialogue systems, conversational agents, and chatbots*. Morgan & Claypool Publishers, 2020.

[24] Matthias Kraus, Nicolas Wagner, Ron Riekenbrauck, and Wolfgang Minker. Improving proactive dialog agents using socially-aware reinforcement learning. In *Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, pages 146–155, 2023.

[25] Prakhar Gupta, Harsh Jhamtani, and Jeffrey P. Bigham. Target-guided dialogue response generation using commonsense and data augmentation, 2022.

[26] Haochen Li, Jonathan Leung, and Zhiqi Shen. Towards goal-oriented prompt engineering for large language models: A survey, 2024.

[27] Jin Wang, JinFei Wang, Shuying Dai, Jiqiang Yu, and Keqin Li. Research on emotionally intelligent dialogue generation based on automatic dialogue system, 2024.

[28] Wenqiang Lei, Yao Zhang, Feifan Song, Hongru Liang, Jiaxin Mao, Jiancheng Lv, Zhenglu Yang, and Tat-Seng Chua. Interacting with non-cooperative user: A new paradigm for proactive dialogue policy, 2022.

[29] Hyunwoo Kim, Youngjae Yu, Liwei Jiang, Ximing Lu, Daniel Khashabi, Gunhee Kim, Yejin Choi, and Maarten Sap. Prosocialdialog: A prosocial backbone for conversational agents, 2022.

[30] Ryosuke Oshima, Seitaro Shinagawa, Hideki Tsunashima, Qi Feng, and Shigeo Morishima. Pointing out human answer mistakes in a goal-oriented visual dialogue, 2023.

[31] Sashank Santhanam and Samira Shaikh. A survey of natural language generation techniques with a focus on dialogue systems - past, present and future directions, 2019.

[32] Tatsuya Kawahara, Koji Inoue, and Divesh Lala. Intelligent conversational android erica applied to attentive listening and job interview, 2021.

[33] Zhuoxuan Jiang, Xian-Ling Mao, Ziming Huang, Jie Ma, and Shaochun Li. Towards end-to-end learning for efficient dialogue agent by modeling looking-ahead ability, 2019.

[34] Jian Wang, Dongding Lin, and Wenjie Li. Target-constrained bidirectional planning for generation of target-oriented proactive dialogue, 2024.

[35] Vladimir Ilievski. Building advanced dialogue managers for goal-oriented dialogue systems, 2018.

[36] Kevin K. Bowden, Jiaqi Wu, Wen Cui, Juraj Juraska, Vrindavan Harrison, Brian Schwarzmann, Nick Santer, and Marilyn Walker. Slugbot: Developing a computational model andframework of a novel dialogue genre, 2019.

[37] A transformer-based user satisfaction prediction for proactive interaction mechanism in dueros.

[38] Jiwei Li. Teaching machines to converse, 2020.

[39] Ting-En Lin, Yuchuan Wu, Fei Huang, Luo Si, Jian Sun, and Yongbin Li. Duplex conversation: Towards human-like interaction in spoken dialogue systems, 2022.

[40] Jinjie Ni, Tom Young, Vlad Pandelea, Fuzhao Xue, and Erik Cambria. Recent advances in deep learning based dialogue systems: A systematic survey. *Artificial intelligence review*, 56(4):3055–3155, 2023.

[41] Jing Yang Lee, Seokhwan Kim, Kartik Mehta, Jiun-Yu Kao, Yu-Hsiang Lin, and Arpit Gupta. Redefining proactivity for information seeking dialogue, 2024.

[42] Sanja Štajner, Seren Yenikent, and Marc Franco-Salvador. Five psycholinguistic characteristics for better interaction with users, 2022.

[43] Haolan Zhan, Yufei Wang, Tao Feng, Yuncheng Hua, Suraj Sharma, Zhuang Li, Lizhen Qu, Zhaleh Semnani Azad, Ingrid Zukerman, and Gholamreza Haffari. Let's negotiate! a survey of negotiation dialogue systems, 2024.

[44] Mengting Wan and Xin Chen. Beyond "how may i help you?": Assisting customer service agents with proactive responses, 2018.

[45] Philip R Cohen. Back to the future for dialogue research: A position paper, 2018.

[46] Yang Deng, Wenqiang Lei, Wai Lam, and Tat-Seng Chua. A survey on proactive dialogue systems: Problems, methods, and prospects, 2023.

[47] Dai Hoang Tran, Quan Z. Sheng, Wei Emma Zhang, Salma Abdalla Hamad, Munazza Zaib, Nguyen H. Tran, Lina Yao, and Nguyen Lu Dang Khoa. Deep conversational recommender systems: A new frontier for goal-oriented dialogue systems, 2020.

[48] Maria João Pereira, Luísa Coheur, Pedro Fialho, and Ricardo Ribeiro. Chatbots' greetings to human-computer communication, 2016.

[49] Samuel Louvan and Bernardo Magnini. Recent neural methods on slot filling and intent classification for task-oriented dialogue systems: A survey, 2020.

[50] Hongru Wang, Lingzhi Wang, Yiming Du, Liang Chen, Jingyan Zhou, Yufei Wang, and Kam-Fai Wong. A survey of the evolution of language model-based dialogue systems, 2023.

[51] Ryuichi Takanobu, Qi Zhu, Jinchao Li, Baolin Peng, Jianfeng Gao, and Minlie Huang. Is your goal-oriented dialog model performing really well? empirical analysis of system-wise evaluation, 2020.

[52] Amit Kumar Chaudhary, Alex J. Lucassen, Ioanna Tsani, and Alberto Testoni. Are current decoding strategies capable of facing the challenges of visual dialogue?, 2022.

[53] Amogh Mannekote. Towards a neural era in dialogue management for collaboration: A literature survey, 2023.

[54] Xinyan Zhao, Bin He, Yasheng Wang, Yitong Li, Fei Mi, Yajiao Liu, Xin Jiang, Qun Liu, and Huanhuan Chen. Unids: A unified dialogue system for chit-chat and task-oriented dialogues, 2021.

[55] Tsung-Hsien Wen, David Vandyke, Nikola Mrksic, Milica Gasic, Lina M. Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. A network-based end-to-end trainable task-oriented dialogue system, 2017.

[56] Xingyu Bruce Liu, Shitao Fang, Weiyan Shi, Chien-Sheng Wu, Takeo Igarashi, and XiangAnthony' Chen. Proactive conversational agents with inner thoughts. *arXiv preprint arXiv:2501.00383*, 2024.

[57] Tian Lan, Xianling Mao, and Heyan Huang. Generative dialog policy for task-oriented dialog systems, 2019.

[58] Julia White, Arushi Raghuvanshi, and Yada Pruksachatkun. Leveraging explicit procedural instructions for data-efficient action prediction, 2023.

[59] Abraham Sanders, Tomek Strzalkowski, Mei Si, Albert Chang, Deepanshu Dey, Jonas Braasch, and Dakuo Wang. Towards a progression-aware autonomous dialogue agent, 2022.

[60] Yutao Zhu, Jian-Yun Nie, Kun Zhou, Pan Du, Hao Jiang, and Zhicheng Dou. Proactive retrieval-based chatbots based on relevant knowledge and goals, 2021.

[61] Jian Wang, Yi Cheng, Dongding Lin, Chak Tou Leong, and Wenjie Li. Target-oriented proactive dialogue systems with personalization: Problem formulation and dataset curation, 2023.

[62] Jing Yang Lee, Seokhwan Kim, Kartik Mehta, Jiun-Yu Kao, Yu-Hsiang Lin, and Arpit Gupta. Redefining proactivity for information seeking dialogue. *arXiv preprint arXiv:2410.15297*, 2024.

[63] Yi-An Lai, Arshit Gupta, and Yi Zhang. Goal-embedded dual hierarchical model for task-oriented dialogue generation, 2019.

[64] Zehao Lin, Shaobo Cui, Guodun Li, Xiaoming Kang, Feng Ji, Fenglin Li, Zhongzhou Zhao, Haiqing Chen, and Yin Zhang. Predict-then-decide: A predictive approach for wait or answer task in dialogue systems, 2021.

[65] Ella Rabinovich, Matan Vetzler, David Boaz, Vineet Kumar, Gaurav Pandey, and Ateret Anaby-Tavor. Gaining insights into unrecognized user utterances in task-oriented dialog systems, 2022.

[66] R. Chulaka Gunasekara, David Nahamoo, Lazaros C. Polymenakos, Jatin Ganhotra, and Kshitij P. Fadnis. Quantized-dialog language model for goal-oriented conversational systems, 2018.

[67] Tom Young. From knowledge augmentation to multi-tasking: Towards human-like dialogue systems, 2022.

[68] Wang Xu, Shuo Wang, Weilin Zhao, Xu Han, Yukun Yan, Yudi Zhang, Zhe Tao, Zhiyuan Liu, and Wanxiang Che. Enabling real-time conversations with minimal training costs, 2024.

[69] Kotaro Shukuri, Ryoma Ishigaki, Jundai Suzuki, Tsubasa Naganuma, Takuma Fujimoto, Daisuke Kawakubo, Masaki Shuzo, and Eisaku Maeda. Meta-control of dialogue systems using large language models, 2023.

[70] Sofia Brenna and Bernardo Magnini. Last utterance proactivity prediction in task-oriented dialogues. In *Proceedings of the Eighth Workshop on Natural Language for Artificial Intelligence (NL4AI 2024) co-located with 23th International Conference of the Italian Association for Artificial Intelligence (AI* IA 2024), CEUR-WS. org*, 2024.

[71] Junkai Zhou, Liang Pang, Huawei Shen, and Xueqi Cheng. Think before you speak: Cultivating communication skills of large language models via inner monologue, 2024.

[72] Yang Deng, Lizi Liao, Liang Chen, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. Prompting and evaluating large language models for proactive dialogues: Clarification, target-guided, and non-collaboration, 2023.

[73] Yash Butala, Siddhant Garg, Pratyay Banerjee, and Amita Misra. Promise: A proactive multi-turn dialogue dataset for information-seeking intent resolution. In *Findings of the Association for Computational Linguistics: EACL 2024*, pages 1774–1789, 2024.

[74] Derek Chen, Howard Chen, Yi Yang, Alex Lin, and Zhou Yu. Action-based conversations dataset: A corpus for building more in-depth task-oriented dialogue systems, 2021.

[75] Chaitanya K. Joshi, Fei Mi, and Boi Faltings. Personalization in goal-oriented dialog, 2017.

[76] Wen-Yu Chang and Yun-Nung Chen. Injecting salesperson's dialogue strategies in large language models with chain-of-thought reasoning, 2024.

[77] Koki Inoue, Shuichiro Ogake, Hayato Kawamura, and Naoki Igo. Dialogue system with humanoid robot, 2022.

[78] Zhao Yan, Nan Duan, Peng Chen, Ming Zhou, Jianshe Zhou, and Zhoujun Li. Building task-oriented dialogue systems for online shopping. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

[79] Yi-Chia Wang, Alexandros Papangelis, Runze Wang, Zhaleh Feizollahi, Gokhan Tur, and Robert Kraut. Can you be more social? injecting politeness and positivity into task-oriented conversational agents, 2020.

[80] Ruifeng Qian, Shijie Li, Mengjiao Bao, Huan Chen, and Yu Che. Toward an optimal selection of dialogue strategies: A target-driven approach for intelligent outbound robots, 2022.

[81] Adi Botea, Christian Muise, Shubham Agarwal, Oznur Alkan, Ondrej Bajgar, Elizabeth Daly, Akihiro Kishimoto, Luis Lastras, Radu Marinescu, Josef Ondrej, Pablo Pedemonte, and Miroslav Vodolan. Generating dialogue agents via automated planning, 2019.

[82] Xueshen Li, Xinlong Hou, Nirupama Ravi, Ziyi Huang, and Yu Gan. A two-stage proactive dialogue generator for efficient clinical information collection using large language model. *arXiv preprint arXiv:2410.03770*, 2024.

[83] Quim Motger, Xavier Franch, and Jordi Marco. Software-based dialogue systems: Survey, taxonomy and challenges, 2024.

[84] Heydar Soudani, Roxana Petcu, Evangelos Kanoulas, and Faegheh Hasibi. A survey on recent advances in conversational data generation, 2024.

[85] Ye Liu, Stefan Ultes, Wolfgang Minker, and Wolfgang Maier. Unified conversational models with system-initiated transitions between chit-chat and task-oriented dialogues, 2023.

[86] Kevin K. Bowden, Shereen Oraby, Amita Misra, Jiaqi Wu, and Stephanie Lukin. Data-driven dialogue systems for social agents, 2017.

[87] Junhua Liu, Yong Keat Tan, Bin Fu, and Kwan Hui Lim. Balancing accuracy and efficiency in multi-turn intent classification for llm-powered dialog systems in production, 2024.

[88] Jessica Van Brummelen, Kevin Weng, Phoebe Lin, and Catherine Yeo. Convo: What does conversational programming need? an exploration of machine learning interface design, 2020.

[89] Tiziano Labruna, Sofia Brenna, Andrea Zaninello, and Bernardo Magnini. Unraveling chatgpt: A critical analysis of ai-generated goal-oriented dialogues and annotations, 2023.

[90] Ruofeng Wen. Turn-level dialog evaluation with dialog-level weak signals for bot-human hybrid customer service systems, 2020.

[91] Eda Okur, Saurav Sahay, and Lama Nachman. Data augmentation with paraphrase generation and entity extraction for multimodal dialogue system, 2022.

[92] Yicheng Zou, Zhihua Liu, Xingwu Hu, and Qi Zhang. Thinking clearly, talking fast: Concept-guided non-autoregressive generation for open-domain dialogue systems. *arXiv preprint arXiv:2109.04084*, 2021.

[93] Sindre André Jacobsen and Anton Ragni. Continuous representations of intents for dialogue systems, 2021.

[94] Antoine Bordes, Y-Lan Boureau, and Jason Weston. Learning end-to-end goal-oriented dialog, 2017.

[95] Jan Deriu, Alvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, and Mark Cieliebak. Survey on evaluation methods for dialogue systems, 2020.

[96] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. A survey on dialogue systems: Recent advances and new frontiers, 2018.

[97] Jan Deriu, Alvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, and Mark Cieliebak. Survey on evaluation methods for dialogue systems. *Artificial Intelligence Review*, 54:755–810, 2021.