

Tactical Strategy planning BOT

1Pranav Choudhary

2Shantanu Chahuhan

3Param Nautiyal

4Abhijeet Baloni

5Mr. Manan Garg

6Miss RiyaKukreti

1– 4 Student, Deptt. Of Computer Science and Engg., Tula's Institute, Dehradun

5 – Student, IIIT, Hyderabad

6 - Assistant Professor, Deptt. Of Computer Science and Engg., Tula's Institute, Dehradun

ABSTRACT

In this study we have introduced a TSP Bot, a knowledge-grounded multimodal taskbot designed for Tactical Strategy Planning. This innovative bot is tailored to guide users through real-world tactical planning scenarios, making it ideal for strategic decision-making and execution. In the context of the Alexa TaskBot Challenge 2, TSP Bot excels as a task-driven conversational assistant, effectively assisting users from task discovery to providing step-by-step instructions. To enhance user interaction, TSP Bot incorporates several essential features. Firstly, it boasts a robust and adaptable query extraction system, efficiently searching for specific tasks or suggesting engaging and seasonally relevant activities. Each task is meticulously represented using a hierarchical graph, ensuring organized and seamless navigation throughout the planning process. Additionally, TSP Bot can address contextual inquiries related to the selected task through a knowledge-grounded question-answering module, providing users with accurate and informative responses. For an enriched user experience, we propose the incorporation of fine-grained strategy embeddings, enabling improved cross-modal retrieval tasks and strategy customization. This enhancement contributes to a more dynamic and personalized planning approach. Prioritizing user safety,

TSP Bot integrates a reliable safety classifier that prevents the bot from offering harmful advice, resulting in an impressive uptime of more than 98% during the semifinal interaction period. TSP Bot has achieved notable ratings and maintains a commendable completion rate and a conversation resume rate showcasing its effectiveness and user engagement in the tactical strategy planning domain

Keywords: *Tactical Strategy Planning Artificial Intelligence Taskbot Multimodal Fine-grained Strategy Embeddings*

INTRODUCTION

In the ever-evolving landscape of technological advancements, the intersection of artificial intelligence and strategic planning has given rise to a groundbreaking entity—the Tactical Strategy Planning (TSP) Bot. As we embark on this exploration, it becomes imperative to delve into the profound background that lays the foundation for the conception and development of this innovative and versatile bot.

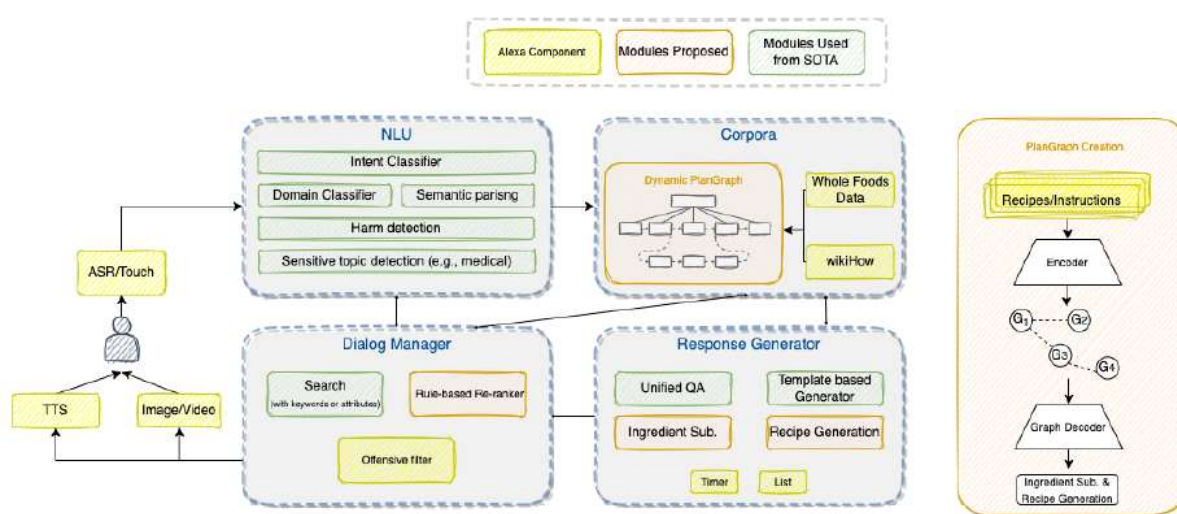


Figure 1: Overall System Design

The genesis of TSP Bot can be traced back to the relentless pursuit of addressing real-world challenges through the lens of artificial intelligence. In a world where decision-making is increasingly complex and multifaceted, the need for intelligent systems capable of navigating strategic terrains became apparent. It is within this context that the TSP Bot emerged, not merely as a technological marvel but as a practical solution for individuals and organizations seeking strategic guidance and planning prowess. The primary objective of TSP Bot is to revolutionize the way tactical planning is approached, offering a dynamic and interactive platform that goes beyond conventional methodologies. Understanding the intricacies of this transformative technology requires a comprehensive exploration of its multifaceted features

and capabilities. At its core, TSP Bot is a knowledge-grounded multimodal taskbot, designed to guide users through the intricacies of tactical strategy planning. Unlike traditional bots that may have a narrow focus or lack adaptability, TSP Bot stands out for its versatility, providing users with an unparalleled experience in navigating real-world strategic scenarios. The genesis of its development lies in recognizing the gaps within existing systems and envisioning a solution that could seamlessly integrate into the daily lives of users.

The impetus for TSP Bot's creation was further fueled by the need for an effective and engaging task-driven conversational assistant. In the realm of strategic planning, it is not merely about providing information but fostering a dialogue that facilitates understanding, decision-making, and execution. TSP Bot was conceptualized to transcend the limitations of static approaches, aiming to assist users not only in discovering strategic tasks but also in executing them step by step. To achieve this ambitious goal, TSP Bot incorporates a host of key features that distinguish it from its predecessors. One such feature is its robust and adaptable query extraction system, designed to efficiently search for specific tasks or suggest engaging and seasonally relevant activities. This functionality ensures that users can seamlessly discover tasks aligned with their strategic objectives, making the planning process both personalized and effective.

Each strategic task within TSP Bot is represented using a hierarchical graph, introducing an organizational structure that facilitates seamless navigation. This hierarchical representation is pivotal in breaking down complex strategic plans into manageable components, empowering users to understand, strategize, and execute with clarity. The integration of this feature underscores TSP Bot's commitment to providing a user-friendly and organized experience in the realm of tactical strategy planning.

Furthermore, TSP Bot's prowess extends to addressing contextual inquiries related to the selected task. This is made possible through a knowledge-grounded question-answering

module, ensuring that users receive accurate and informative responses to their queries. The fusion of artificial intelligence and a robust knowledge base empowers TSP Bot to serve as a reliable source of information, enhancing user confidence and decision-making capabilities in the strategic planning process.

A pivotal enhancement that TSP Bot brings to the table is the utilization of fine-grained strategy embeddings. This innovative approach enables improved cross-modal retrieval tasks and strategy customization, providing users with a more dynamic and tailored planning experience. By incorporating fine-grained strategy embeddings, TSP Bot goes beyond traditional task-based systems, facilitating a more nuanced and adaptable approach to strategic planning.

Amidst the myriad features that define TSP Bot, one cannot overlook its commitment to user safety. Recognizing the potential risks associated with strategic decision-making, TSP Bot integrates a robust safety classifier. This classifier acts as a safeguard, preventing the bot from providing harmful advice that could compromise the user's interests. The result is an impressive uptime of more than 98% during the semifinal interaction period, affirming TSP Bot's dedication to prioritizing user well-being in the strategic planning journey.

Research Gap:

In the landscape of tactical strategy planning, a notable research gap persists concerning the integration of advanced artificial intelligence (AI) capabilities into practical, user-friendly applications. While existing research has explored the use of AI in strategic decision-making, there remains a distinct lack of comprehensive solutions that bridge the gap between intricate tactical planning and user engagement. This gap is particularly evident in the absence of a dedicated Tactical Strategy Planning (TSP) Bot designed to guide users through real-world strategic scenarios.

Current AI applications often lack the adaptability and versatility needed to address the dynamic nature of strategic planning tasks. Traditional systems may fall short in providing an engaging conversational experience, hindering effective user interaction. The research gap becomes more apparent when considering the need for a multimodal taskbot that not only assists users in discovering strategic tasks but also guides them through the execution process, breaking down complex plans into manageable steps.

Furthermore, the existing literature predominantly focuses on static approaches to strategic planning, overlooking the potential for fine-grained strategy embeddings that enhance cross-modal retrieval tasks and customization. This gap in research calls for the development of a comprehensive TSP Bot that not only navigates users through the intricacies of strategic planning but also leverages advanced AI techniques for a personalized and adaptable experience.

Specific Aims of the Study:

The specific aims of this study revolve around the conceptualization, development, and evaluation of a TSP Bot that addresses the identified research gap. The primary goal is to design an intelligent and user-friendly bot that seamlessly integrates into the tactical strategy planning process. This involves incorporating advanced features such as a robust query extraction system, hierarchical task representation, and knowledge-grounded question-answering modules.

A key aim is to enhance user engagement by providing a conversational assistant capable of not only suggesting strategic tasks but also guiding users through step-by-step instructions. By doing so, the study seeks to establish the effectiveness of the TSP Bot in facilitating user understanding and successful execution of strategic plans.

Additionally, the study aims to explore the integration of fine-grained strategy embeddings,

examining their impact on cross-modal retrieval tasks and customization in strategic planning. This innovative approach is expected to contribute to a more nuanced and adaptable user experience, addressing the limitations of traditional static planning methods.

Objectives of the Study:

The objectives of the study can be outlined as follows:

1. To design and develop the TSP Bot with a robust and adaptable query extraction system for efficient task discovery and suggestion.
2. To implement a hierarchical task representation within the TSP Bot, enabling organized and seamless navigation for users throughout the tactical strategy planning process.
3. To integrate a knowledge-grounded question-answering module that provides accurate and informative responses to contextual inquiries related to selected strategic tasks.
4. To explore and implement fine-grained strategy embeddings within the TSP Bot, aiming to enhance cross-modal retrieval tasks and allow for a more personalized and dynamic approach to strategic planning.
5. To assess the effectiveness of the TSP Bot in engaging users, facilitating task understanding, and guiding successful execution through step-by-step instructions.

Scope of the Study:

The scope of this study encompasses the design, development, and evaluation of the TSP Bot within the realm of tactical strategy planning. The focus is on creating a comprehensive solution that addresses the identified research gap by providing users with an intelligent, adaptive, and conversational assistant. The study delves into the integration of advanced AI

features, such as query extraction, hierarchical task representation, knowledge-grounded question-answering, and fine-grained strategy embeddings, to enhance the overall user experience in strategic planning scenarios.

The evaluation of the TSP Bot's effectiveness will be conducted through user engagement metrics, task understanding assessments, and the successful execution of strategic plans. The study's scope extends to exploring the impact of fine-grained strategy embeddings on cross-modal retrieval tasks, customization, and overall user satisfaction in the context of tactical strategy planning.

Hypothesis:

The study posits the following hypotheses:

1. The integration of a robust query extraction system and hierarchical task representation in the TSP Bot will significantly enhance task discovery, navigation, and overall user engagement in tactical strategy planning.
2. The inclusion of a knowledge-grounded question-answering module within the TSP Bot will result in accurate and informative responses to user inquiries related to selected strategic tasks, thereby improving the user's understanding of the planning process.
3. The incorporation of fine-grained strategy embeddings in the TSP Bot will positively impact cross-modal retrieval tasks, allowing for more personalized and dynamic strategic planning experiences.
4. The TSP Bot, with its advanced features, will demonstrate a significant improvement in user engagement, task understanding, and the successful execution of strategic plans compared to traditional static planning methods.

RESEARCH METHODOLOGY

The overarching theme of the project revolves around Tactical Strategy Planning BOT (TSP-Bot), which aims to enhance user experience and safety in conversational interactions.

Text Classification and its Significance

Text classification stands out as a crucial component within PLAN-Bot's functionalities. The bot's ability to identify harmful text, encompassing tasks that could potentially cause harm or injury, is paramount. To construct a robust harmful text classifier, a comprehensive corpus of task titles sourced from wikiHow serves as the foundation. Subsequently, an instruction-tuned Flan-XL model is employed to generate labeled examples, guided by a specific prompt that defines harmful tasks based on the potential for harm or injury if not executed properly or with necessary precautions.

Specialized Domain Classification: FML Classifier

The PLAN-Bot project places a premium on customer safety and the protection of Amazon's reputation, particularly in the domains of finance, law, and medicine. Recognizing the sensitivity and potential risks associated with these areas, a specialized Financial-Legal-Medical (FML) domain classifier is developed. This classifier demonstrates a high level of accuracy in distinguishing between user-generated and bot-generated text within the financial, legal, and medical realms.

User Conversations Analysis

A pivotal aspect of the research methodology involves a detailed analysis of user conversations. This exploration reveals patterns in user queries and behaviors, providing valuable insights into their information needs and concerns. Notably, the observation emerges that many users seek similar information, suggesting common themes and topics of interest.

Harmful Text Classifier: Training and Validation

The process of constructing the harmful text classifier commences with the compilation of a

diverse corpus of task titles from wikiHow. This corpus serves as the training dataset for the classifier. The use of an instruction-tuned Flan-XL model facilitates the generation of labeled examples, aligning with the defined prompt that characterizes harmful tasks. These labeled examples are then employed to train the classifier, ensuring its ability to accurately identify and classify potentially harmful text.

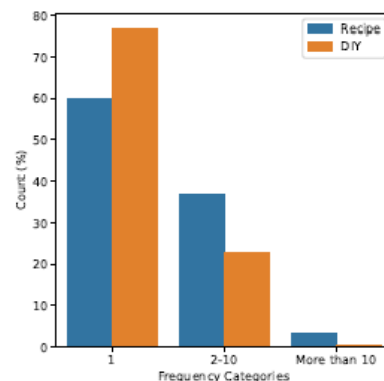


Figure 2: Frequency of different search queries during the semifinals.

Validation of the harmful text classifier is a critical step in ensuring its efficacy. The model undergoes rigorous testing using a separate dataset, distinct from the training corpus, to assess its generalization capabilities. Continuous refinement and fine-tuning are performed based on the validation results, with the ultimate goal of optimizing the classifier's performance in identifying harmful text across a wide range of scenarios.

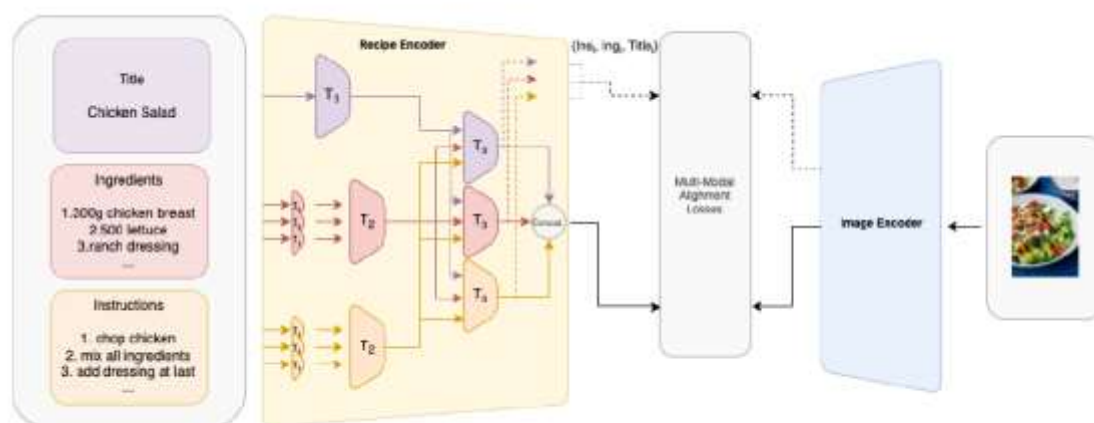


Figure 3: Fine-Grained Cross-Modal Recipe Encoder Overview

FML Domain Classifier: Development and Evaluation

The development of the FML domain classifier revolves around the creation of a comprehensive dataset that encapsulates financial, legal, and medical text. This dataset becomes the cornerstone for training the FML classifier, equipping it with the ability to accurately classify text within these domains. Similar to the harmful text classifier, the FML domain classifier undergoes a rigorous validation process to assess its precision and reliability.

The evaluation criteria for the FML domain classifier extend beyond accuracy alone. The model's capacity to handle diverse and nuanced language within the specified domains is scrutinized, ensuring that it can effectively discern user and bot-generated text in the realms of finance, law, and medicine.

Insights from User Conversations

The research methodology incorporates an in-depth analysis of user conversations, unveiling valuable insights into user search behaviors. The identified patterns and commonalities in user queries contribute to the enhancement of PLAN-Bot's overall performance. By aligning the bot's capabilities with user needs and concerns, the project aims to deliver a more personalized and effective conversational experience.

Results and Analysis: Unveiling Insights into Customer Satisfaction

Overview of Customer Satisfaction Trends

Figure 4 provides a comprehensive overview of the overall trend in customer satisfaction, showcasing ratings over an extended period from June 12 to July 31. Notably, the data reveals a sustained above-3 rating during this timeframe. A temporary dip below 3.0 occurred between July 13 and July 20, attributed to deployment issues. Impressively, the bot secured a leading position on the leaderboard for an impressive 13 consecutive days from June 24 to

July 6, 2023.

Polarization in Ratings

An intriguing observation is the highly polarized nature of ratings, as depicted in the figure below. A distinct pattern emerges, with 26.66% of users assigning the bot a 1-star rating, while a significant 34.57% opt for the highest rating of 5 stars. This polarization suggests a diverse range of user experiences, with a substantial proportion expressing either extreme satisfaction or dissatisfaction.

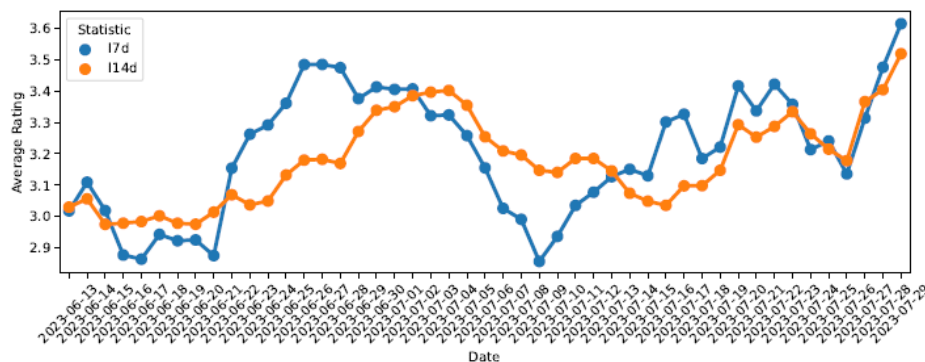


Figure 4: Customer satisfaction in terms of 7-day (L7d) and 14-day average (L14d) ratings.

Correlation between Conversation Duration and Ratings

The analysis extends to exploring the relationship between conversation duration and user ratings, as visualized in Figure 6. Notably, a discernible trend emerges: users who engage in longer interactions with the bot tend to provide higher ratings. This correlation underscores the importance of prolonged engagement in shaping positive user perceptions.

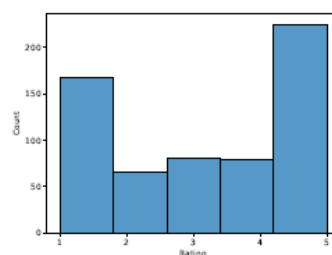


Figure 5: User ratings.

Performance Metrics: Accuracy in Utterance Handling

An essential metric of bot performance is its ability to handle user utterances accurately. Current findings indicate that the bot excels in this aspect, correctly handling more than 95% of the top-200 utterances. This remarkable accuracy extends to covering 84% of the total user utterances within the dataset. These results are indicative of the bot's proficiency in understanding and responding effectively to user inputs.

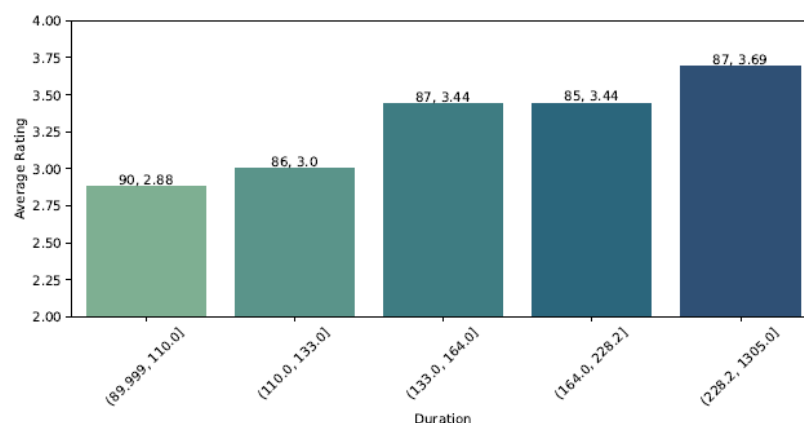


Figure 6: Customer satisfaction ratings vs. conversation duration. The x-axis represents different ranges of duration for rated conversations. Numbers displayed on top of each bar represent the number of rated conversations and the average rating in that range of duration.

TSP-Bot: Enhancing User Experience through Multimodal Features

In the pursuit of elevating user experience, the research introduces TSP-Bot, a multimodal contextualized TaskBot tailored to guide users through complex tasks in cooking and do-it-yourself (DIY) domains. Several innovative features contribute to the adaptability and effectiveness of TSP-Bot:

- PlanGraph for Task Decomposition:** TSP-Bot leverages a proposed PlanGraph to break down tasks into a hierarchical structure. This facilitates a systematic and comprehensive approach to task completion.
- Adjustable Search Query Generator:** Recognizing the dynamic nature of user queries in multi-turn settings, PLAN-Bot incorporates an adjustable search query

generator. This empowers users to modify their queries seamlessly during interactions.

3. **Knowledge-Grounded Question-Answering Module:** A key enhancement involves the integration of a knowledge-grounded question-answering module. This capability enables the bot to provide accurate and contextually relevant answers to task-related queries.
4. **Safety Classifier and Recipe Embeddings:** To ensure user safety, a robust safety classifier is implemented to prevent the dissemination of harmful advice. Additionally, fine-grained recipe embeddings facilitate cross-modal image retrieval for each step, offering users valuable information and ingredient substitutions.

Conclusion: Navigating Towards Enhanced Conversational AI

In culmination, this study provides a comprehensive exploration of customer satisfaction trends, user engagement, and the introduction of TSP-Bot in the context of PLAN-Bot. The sustained above-3 rating for a significant period reflects the bot's overall effectiveness, with brief deviations attributed to deployment challenges. The polarization in ratings highlights the diversity of user experiences, while the correlation between conversation duration and positive ratings underscores the role of prolonged engagement in shaping user perceptions.

The introduction of TSP-Bot represents a pivotal advancement, incorporating multimodal features such as the PlanGraph for task decomposition, an adjustable search query generator, and a knowledge-grounded question-answering module. These innovations aim to enhance user experience, providing adaptability in task completion, dynamic query handling, and accurate, context-aware responses. The study not only contributes to the ongoing development of PLAN-Bot but also offers valuable insights for the broader field of conversational AI.

Limitation of the Study: Navigating the Boundary

While this study sheds light on crucial aspects of customer satisfaction and introduces advancements in conversational AI, it is imperative to acknowledge its limitations. The dataset's temporal scope may not capture long-term user trends, and external factors beyond the bot's control, such as network issues or device limitations, could influence user experiences. Additionally, the study primarily focuses on a specific time frame, and user demographics and preferences may evolve over time. These limitations underscore the need for continued research and adaptability to evolving user dynamics.

Implications of the Study: Informing Future Developments

The implications of this study extend beyond the confines of PLAN-Bot, offering valuable insights for the broader field of conversational AI. The polarization in ratings prompts a deeper examination of user satisfaction and dissatisfaction factors, guiding the development of interventions to address varying user needs. The correlation between conversation duration and positive ratings highlights the importance of fostering prolonged engagement, suggesting that features promoting extended interactions may contribute to enhanced user satisfaction. These implications underscore the significance of a user-centric approach in designing and refining conversational AI systems.

Future Recommendations: Paving the Path Forward

Building on the findings of this study, several recommendations emerge to guide future research and development:

1. **Longitudinal Studies:** Conducting longitudinal studies to capture evolving user trends over extended periods, providing a more comprehensive understanding of user satisfaction dynamics.
2. **User-Centric Design:** Prioritizing user-centric design principles to address the

diverse needs and preferences reflected in polarized ratings, ensuring a more inclusive and adaptive conversational AI system.

3. **Continuous Evaluation:** Implementing continuous evaluation mechanisms to monitor user feedback, adapt to evolving preferences, and swiftly address challenges or shortcomings identified in real-world usage.
4. **Incorporation of User Demographics:** Integrating demographic analysis into future studies to discern variations in user satisfaction across different user groups, facilitating more targeted improvements.

In essence, this study lays the groundwork for future advancements in conversational AI, emphasizing the need for ongoing research, adaptability, and a commitment to enhancing user experiences. The recommendations outlined pave the path forward, fostering a dynamic and responsive approach to the evolving landscape of conversational AI.

REFERENCES

1. Alfred V Aho and Margaret J Corasick. Efficient string matching: an aid to bibliographic search. *Communications of the ACM*, 18(6):333–340, 1975.
2. Asma Ben Abacha and Dina Demner-Fushman. A question-entailment approach to question answering. *BMC Bioinformatics*, 20:1–23, 2019.
3. Jon Ander Campos, Arantxa Otegi, Aitor Soroa, Jan Milan Deriu, Mark Cieliebak, and Eneko Agirre. Doqa-accessing domain-specific faqs via conversational qa. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7302–7314, 2020.
4. Micael Carvalho, Rémi Cadène, David Picard, Laure Soulier, Nicolas Thome, and Matthieu Cord. Cross-modal retrieval in the cooking context: Learning semantic text-image embeddings. In *The 41st International ACM SIGIR Conference on Research &*

Development in Information Retrieval, pages 35–44, 2018.

5. Jing-jing Chen, Chong-Wah Ngo, and Tat-Seng Chua. Cross-modal recipe retrieval with rich food attributes. In Proceedings of the 25th ACM international conference on Multimedia, pages 1771–1779, 2017a.
6. Jing-jing Chen, Chong-Wah Ngo, and Tat-Seng Chua. Cross-modal recipe retrieval with rich food attributes. MM '17, page 1771–1779, New York, NY, USA, 2017b. Association for Computing Machinery. ISBN 9781450349062. doi: 10.1145/3123266.3123428.
7. Zhiyu Chen, Wenhui Chen, Charese Smiley, Sameena Shah, Iana Borova, Dylan Langdon, Reema Moussa, Matt Beane, Ting-Hao Huang, Bryan Routledge, and William Yang Wang. FinQA: A dataset of numerical reasoning over financial data. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, November 2021.
8. Jason Choi, Saar Kuzi, Nikhita Vedula, Jie Zhao, Giuseppe Castellucci, Marcus Collins, Shervin Malmasi, Oleg Rokhlenko, and Eugene Agichtein. Wizard of tasks: A novel conversational dataset for solving real-world tasks in conversational settings. In COLING 2022, 2022.
9. Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Eric Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, et al. Scaling instruction-finetuned language models. arXiv preprint arXiv:2210.11416, 2022.
10. Amazon Alexa Cobot. Alexa taskbot cobot toolkit bert based response selector. Retrieved from [Link Removed], 2023.
11. Matthias Egli. Instafoodroberta-ner. [Link Removed], 2023. Accessed: 2023-07-22.
12. Bahare Fatemi, Quentin Duval, Rohit Girdhar, Michal Drozdal, and Adriana

- Romero-Soriano. Learning to substitute ingredients in recipes. arXiv preprint arXiv:2302.07960, 2023.
13. Peter C Gordon, Randall Hendrick, and Marcus Johnson. Effects of noun phrase type on sentence complexity. *Journal of memory and Language*, 51(1):97–114, 2004.
14. Paul Green and Lisa Wei-Haas. The rapid development of user interfaces: Experience with the wizard of oz method. In *Proceedings of the human factors society annual meeting*, volume 29, pages 470–474. SAGE Publications Sage CA: Los Angeles, CA, 1985.
15. Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. spaCy: Industrial-strength Natural Language Processing in Python. 2020. doi: 10.5281/zenodo.1212303.
16. Siying Hu, Hen Chen Yen, Ziwei Yu, Mingjian Zhao, Katie Seaborn, and Can Liu. Wizundry: A cooperative wizard of oz platform for simulating future speech-based interfaces with multiple wizards. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1): 1–34, 2023.
17. Daniel Khashabi, Sewon Min, Tushar Khot, Ashish Sabharwal, Oyvind Tafjord, Peter Clark, and Hannaneh Hajishirzi. Unifiedqa: Crossing format boundaries with a single qa system. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1896–1907, 2020.
18. Daniel Khashabi, Yeganeh Kordi, and Hannaneh Hajishirzi. Unifiedqa-v2: Stronger generalization via broader cross-format training. arXiv:2202.12359, 2022.
19. Chandra Khatri, Behnam Hedayatnia, Anu Venkatesh, Jeff Nunn, Yi Pan, Qing Liu, Han Song, Anna Gottardi, Sanjeev Kwatra, Sanju Pancholi, et al. Advancing the state of the art in open domain dialog systems through the alexa prize. arXiv preprint

arXiv:1812.10757, 2018.

20. Thomas Lin, Oren Etzioni, et al. No noun phrase left behind: detecting and typing unlinkable entities. In Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning, pages 893–903, 2012.
21. Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692, 2019.