Abstract

This research explores the innovative integration of commonsense knowledge (CSK) within AI systems, with a particular focus on domestic robotics. At the heart of this study is the Robo-CSK-Organizer, a groundbreaking system that utilizes a classical knowledge base, namely ConceptNet, to enhance robotic decision-making through sophisticated object organization and classification. This system is contrasted with a ChatGPT-based organizer, examining their performance in terms of ambiguity resolution, consistency in object placement, adaptability to task classifications, and crucially, in explainability, a key aspect of XAI (Explainable AI). Through a combination of controlled experiments, quantitative and qualitative analysis, the study demonstrates that the Robo-CSK-Organizer not only classifies objects accurately, but also surpasses related systems such as the ChatGPT organizer in XAI. It dynamically applies CSK, thereby offering clearer and more understandable decision-making processes. These advancements are significant for the field of multipurpose robotics, suggesting enhanced user experience, increased trust, improved error detection and correction, as well as better comprehension of robotic actions by users. This thesis is distinguished by its comprehensive comparative analysis of CSK applications in AI, particularly highlighting the importance of XAI. It introduces an innovative framework for evaluating how AI systems resolve ambiguity and make decision-making processes transparent and interpretable. The study addresses the challenges of integrating CSK into robotics, focusing on ambiguity, consistency, task relevance, and explainability in object classification. It asserts that the Robo-CSK-Organizer, with its specialized deployment of a knowledge base (ConceptNet), not only achieves higher accuracy and consistency in object classification and placement but also provides superior explainability compared to a base case organizer that uses ChatGPT. The significance of this study lies in its

contribution to AI research, offering empirical evidence on the strengths and limitations of CSK integration, especially in the context of XAI. It advances the conversation around AI transparency and demonstrates the practical application of CSK in multipurpose robotics. By illustrating how semantic task planning and CSK can be effectively operationalized in service robots with an emphasis on explainability, this research bridges an essential gap between theoretical concepts and their practical implementation in AI systems.

MONTCLAIR STATE UNIVERSITY

Infusing Commonsense via Knowledge Bases in Multipurpose Robotic Task Organization

By

Rafael Omar Hidalgo

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Computer Science

January 2024

College of Science and Mathematics

Montclair School of Computing

Thesis Committee:

Dr. Aparna Varde
Thesis Sponsor

Dr. Rui Li
Committee Member

Dr. Weitian Wang
Committee Member

Infusing Commonsense via Knowledge Bases in Multipurpose Robotic Task Organization

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of **Computer Science**

By

Rafael Hidalgo

Montclair State University

Montclair, NJ

2024

Dedication

To my parents,

I want it to be known that I deeply appreciate the sacrifices and hardships you endured to support me on my life journey. As immigrants who left their homelands to forge a new life here, you kindled a flame of opportunity. I am committed to nurturing that flame, ensuring it burns brighter with each passing day, and for future generations to come.

To my fiancée, Ashley Campusano,

I extend my heartfelt thanks for standing by my side throughout my educational journey. Your unwavering support and the goals we've shaped together have not only enriched my life but have also fanned the flames of ambition that our parents ignited. Your love and dedication have been a constant source of inspiration. Thank You.

Acknowledgments

I extend my sincere appreciation to my thesis advisor, Dr. Aparna Varde, not only for her unwavering support and guidance but also for her belief in my potential. Under her tutelage, I have achieved milestones that I once thought were beyond my reach.

I am grateful to Dr. Rui Li and Dr. Weitian Wang for their invaluable contributions as members of my thesis committee, providing their expertise and insights.

Special thanks go to my research partner, Jesse Parron, whose extensive knowledge in robotics has been instrumental in making my research possible. Much like in Shōnen, Our collaboration has demonstrated that with determination and teamwork, any challenge can be overcome.

I also want to acknowledge the School of Computer Science at Montclair State University for providing me with the opportunity to pursue my graduate studies.

I am indebted to my Graduate Program Coordinator, Dr. Boxiang Dong, and the Chair of the Computer Science Department, Dr. Constantine Coutras, for their unwavering support and guidance throughout my academic journey.

We acknowledge NSF MRI grant "MRI: Acquisition of a Multimodal Collaborative Robot System (MCROS) to Support Cross-Disciplinary Human-Centered Research and Education at Montclair State Universi-ty", Award# 2117308. Our experiments are conducted in the CRoSS (Col-laborative Robotics and Smart Systems) Lab at MSU, of which Dr. Weitian Wang is the Director. We also acknowledge the support of CESAC (Clean Energy and Sustainability Analytics Center) at MSU, of which Dr. Aparna Varde is an Associate Director. We thank the School of Computing at MSU for their support. MSU is a designated HSI (Hispanic Serving Institution) and a DRU-2 (Doctoral Research University with high research status).

Table of Contents

List of Acronyms

R-CNN: Region-Based Convolutional Neural Networks

LSTM: Long Short-Term Memory

NLP: Natural Language Processing

HRI: Human-Robot Interaction

XAI: Explainable Artificial Intelligence

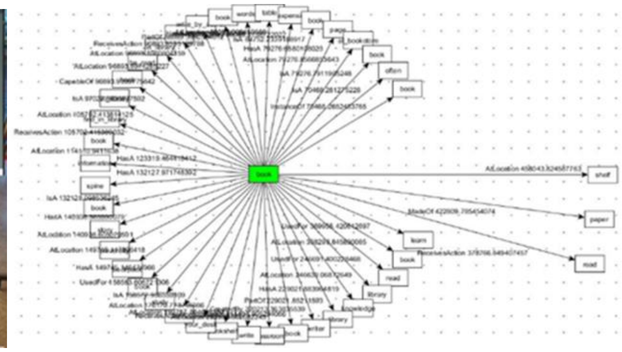
ROS: Robot Operating System

BLIP: Bootstrapping Language-Image Pre-training

DETIC: Detector Incorporating Image-level Classification

AI: Artificial Intelligence

List of Figures



Chapter 1: Introduction

1.1 Background and Motivation

In the evolving field of artificial intelligence (AI), the integration of commonsense knowledge (CSK) into robotic systems is crucial for enhancing their decision-making capabilities. The prevalence of domestic robots is on the rise, and with it, the imperative to navigate the complexities of human environments, which are often laden with implicit rules and nuanced contexts. Traditional AI systems frequently struggle with tasks that require a level of contextual awareness akin to human understanding, such as distinguishing between a child's toy and a pet's toy, or recognizing that a cup left on a bedside table overnight should not be relocated to the kitchen. These scenarios underscore the limitations of current AI systems and the necessity for CSK to bridge the gap between robotic perception and human-like reasoning.

The Robo-CSK-Organizer, the focus of this thesis, is a system designed to imbue robots with the ability to organize and classify objects using CSK, drawing from the rich semantic network of the ConceptNet knowledge base. It is engineered to make transparent decisions, such as determining the most logical location for a pear, whether in the kitchen or the garden, based on contextual cues. This capability is pivotal in addressing the "opaque-box" issue prevalent in AI, where the rationale behind decisions often remains obscured. The challenges of explainability in AI, as articulated in the paper "Solving the Black Box Problem: A Normative

Framework for Explainable Artificial Intelligence" (Reference), highlight the growing demand for systems that are not only intelligent but also interpretable and trustworthy.

This study adopts a rigorous methodological approach to compare the Robo-CSK-Organizer with a ChatGPT-based implementation, employing controlled experiments and quantitative analysis to evaluate performance. The Robo-CSK-Organizer demonstrates not only superior accuracy in object classification tasks but also introduces innovative techniques for the dynamic application of CSK, considering spatial, temporal, and functional aspects of commonsense reasoning.

The practical implications of this research are profound, with the potential to enhance user experience significantly and foster increased trust in domestic robots. The societal impact of such advancements could lead to more autonomous and helpful domestic aides, while the economic implications might include reduced costs associated with human oversight and increased efficiency in household management. By situating the Robo-CSK-Organizer within the current academic discourse, as evidenced by "Large Language Models as Commonsense Knowledge for Large-Scale Task Planning" [1], this study contributes to a deeper understanding of CSK integration in AI. Moreover, the real-world applications of CSK in AI are exemplified in "Semantic Task Planning for Service Robots in Open World," which presents solutions to the challenges that the Robo-CSK-Organizer is designed to overcome [2].

1.2 Problem Statement

The challenge of integrating commonsense knowledge (CSK) into robotic systems for domestic environments is multifaceted, encompassing issues of inherent ambiguity, consistency, task relevance, and explainability in object classification. For instance, the classification of a pear, which could be contextually placed in either the kitchen or the garden, presents a dilemma that is emblematic of the broader issue of ambiguity in object categorization. This is not merely a theoretical concern but a practical one, as evidenced by the difficulties faced by assistive robots in discerning the stability and occlusion of objects in various configurations, as highlighted in the paper "Towards combining commonsense reasoning and knowledge acquisition to guide deep learning" [3]. The paper discusses the challenges of providing labeled training examples for all possible arrangements of objects, a task that is both labor-intensive and computationally demanding.

Moreover, the expectation that a robot consistently places an object, such as a remote control, in the living room regardless of contextual variations is crucial for maintaining user trust. This consistency must be achieved in the face of incomplete domain knowledge and associated uncertainties, as the robot must reason with both qualitative and quantitative descriptions of its environment. The paper provides a compelling example of an assistive robot tasked with clearing away toys in different rooms, where the robot must navigate through incomplete and qualitative descriptions of domain knowledge, such as default statements about the stability of object configurations.

Additionally, the robot's ability to prioritize tasks, such as deciding whether to engage in gardening or culinary activities, can significantly influence its classification decisions. This adaptability to task relevance is critical, especially when considering the robot's decision-making process in the presence of probabilistic estimates of uncertainty in sensing and navigation, as well as the limitations in human participants' ability to provide comprehensive feedback.

Finally, a significant aspect of these challenges is the issue of explainability in AI systems, particularly pertinent in the context of domestic robotics. The Robo-CSK-Organizer, at the center of this research, is designed to not only classify objects accurately but also to surpass related systems like the ChatGPT organizer in terms of explainability, a key aspect of XAI (Explainable AI). This system demonstrates the dynamic application of CSK, offering clearer and more understandable decision-making processes. The ability to elucidate the rationale behind decisions made by AI systems is crucial for ensuring that these systems are not only intelligent but also transparent and trustworthy, allowing users to comprehend and have confidence in the decisions made by their robotic assistants.

By addressing these specific challenges, the Robo-CSK-Organizer aims to enhance the decision-making capabilities of domestic robots, enabling them to navigate the complexities of human environments with greater reliability, precision, and transparency.

1.3 Objectives of the Study

- **Quantify the Efficacy of the Robo-CSK-Organizer:** Investigate the efficacy of the Robo-CSK-Organizer in resolving ambiguity in object classification by measuring its accuracy rate in correctly categorizing objects across diverse contexts. The system's decisions will be compared against a set of NLP methods: FastText, GloVe, and Word2Vec. This standard will serve as an objective benchmark for both the Robo-CSK-Organizer and the ChatGPT Organizer.

- **Measure the Consistency of Object Placement:** Evaluate the consistency of the Robo-CSK-Organizer and ChatGPT Organizer in object placement by analyzing the variance in placement decisions across repeated trials in identical scenarios. Statistical methods will be employed to determine variance, with a lower variance indicating a higher level of consistency, which is crucial for user predictability and trust in the system's behavior.

- **Assess Adaptability to Task-Relevant Classifications:** Assess the adaptability of both systems to task-relevant classifications by implementing a series of dynamic scenarios where the priority of tasks changes. The systems' ability to adjust object classification decisions in response to these changes will be measured, focusing on the speed and accuracy of adaptation as indicators of system flexibility. Specific metrics for adaptation will include the time taken to adjust to new priorities and the rate of correct classifications post-adaptation.

- **Evaluating transparency in the decision-making process.** An important aspect of this study is to impartially compare the transparency of the decision-making processes of both the Robo-CSK-Organizer and the ChatGPT Organizer. This evaluation will objectively assess how understandable and traceable each system's reasoning and decision pathways are. It is hypothesized that the decision-making process of the Robo-CSK-Organizer may exhibit different levels of transparency due to its use of the ConceptNet knowledge base. However, this will be systematically evaluated against the ChatGPT Organizer to determine if there is a significant difference in the clarity of reasoning provided by each system.

1.4 Significance of the Study

The findings of this study are poised to make a substantial impact on the field of general and domestic robotics by providing a deeper understanding of how commonsense knowledge (CSK) can be integrated into artificial intelligence (AI) systems for improved functionality in real-world settings. Specifically, the study's insights could lead to:

Enhanced Research Directions: The study's comparative analysis is set to redefine research directions by uncovering the nuances of CSK application in AI, offering a blueprint for developing versatile CSK frameworks. These frameworks could be adapted for a range of AI platforms, catalyzing a wave of innovation in the field. Additionally, by dissecting the capabilities of ConceptNet within AI systems, this research could pave the way for enhanced knowledge bases tailored to the intricacies of practical AI tasks.

Technology Development: The practical demonstration of CSK in domestic robots within this study is expected to revolutionize AI system design, fostering the creation of more empathetic and anticipatory AI. The anticipated outcomes include AI systems with an unprecedented ability to comprehend and anticipate human needs, setting a new benchmark for human-AI synergy.

Industry Practices: The findings are anticipated to be a catalyst for change in industry practices, underscoring the critical role of CSK in elevating AI system dependability. This could instigate a shift towards new AI development standards, particularly for consumer robotics, ensuring these systems are adept at navigating the unpredictability of human living spaces.

Policy and Ethical Considerations: By spotlighting the importance of explainability and trust in AI, the study is expected to have a significant influence on the formulation of policies and ethical standards that govern AI development. The goal is to ensure AI systems not only perform tasks but also align with societal values and operate with transparency and accountability.

Educational Implications: The research methodologies and outcomes have the potential to transform AI and robotics education, equipping future innovators with the knowledge to integrate CSK into AI. This could lead to a new era of AI practitioners who prioritize the seamless integration of AI into human environments.

1.5 Thesis Contributions

Providing a Comparative Analysis of Robo-CSK-Organizer and ChatGPT Organizer: While large language models have been explored as a source of commonsense knowledge for task planning, as discussed in "Large Language Models as Commonsense Knowledge for Large-Scale Task Planning", there is a lack of comparative studies that evaluate their effectiveness against specialized systems like the Robo-CSK-Organizer [1]. This thesis fills this gap by offering a direct comparison, providing empirical evidence on the strengths and limitations of each approach in domestic robotics.

Introducing a Novel Framework for Evaluating Ambiguity Resolution in AI: The challenge of explainability in AI, particularly in the context of decision-making, is a central concern in "Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence" [4]. This thesis extends the discourse by proposing a novel framework that not only evaluates the transparency of AI decisions but also their ability to resolve ambiguity in object classification, a critical aspect that has not been sufficiently addressed in existing models.

Demonstrating the Practical Implications of CSK in Multipurpose Robots: The integration of commonsense reasoning and knowledge acquisition is pivotal for enhancing AI performance, as suggested in "Towards combining commonsense reasoning and knowledge acquisition to guide deep learning" [3]. This thesis goes a step further by demonstrating how these theoretical concepts can be applied in practice, using the Robo-CSK-Organizer as a case study. It showcases the real-world implications of CSK, particularly in next-generation multipurpose domestic robots, which is a direct application of the principles outlined in

"Semantic Task Planning for Service Robots in Open World" [2]. By doing so, it provides a tangible example of how semantic task planning and CSK can be operationalized in service robots, addressing a gap in the translation of theory to practice.

1.6 Research Questions and Hypotheses

The study is guided by the following research questions and hypotheses:

Research Question 1: How does the Robo-CSK-Organizer resolve the ambiguity in object classification compared to the ChatGPT Organizer?

Hypothesis 1: The Robo-CSK-Organizer is hypothesized to demonstrate a similar accuracy rate in resolving ambiguity in object classification compared to the ChatGPT Organizer, potentially due to its specialized use of the ConceptNet knowledge base and semantic reasoning algorithms.

Research Question 2: Is there a significant difference in the consistency of object placement between the two systems?

Hypothesis 2: It is hypothesized that the Robo-CSK-Organizer will exhibit greater consistency in object placement across various contexts than the ChatGPT Organizer. This may be attributed to its algorithms that are specifically designed to maintain context-specific object relations, leveraging the structure and connections within ConceptNet.

Research Question 3: How do task-relevant priorities affect the classification decisions of each system?

Hypothesis 3: When presented with changing task-relevant priorities, both systems are expected to modify their classification decisions. However, the Robo-CSK-Organizer is hypothesized to adapt more effectively, possibly due to its integrated CSK framework that accounts for temporal and functional aspects of commonsense reasoning.

Research Question 4: How do the Robo-CSK-Organizer and the ChatGPT Organizer differ in the transparency of their decision-making processes?

Hypothesis 4: The Robo-CSK-Organizer is expected to provide greater transparency in its decision-making process compared to the ChatGPT Organizer, owing to its explicit use of the ConceptNet knowledge base, which may facilitate more understandable and traceable reasoning paths for its decisions.

1.7 Scope and Delimitations

This study is specifically focused on the comparative analysis of two AI systems: the Robo-CSK-Organizer, which utilizes the ConceptNet knowledge base alongside semantic reasoning algorithms to facilitate commonsense understanding in object classification and placement; and the ChatGPT Organizer, which leverages a language model-based approach for decision-making processes.

The scope of this research encompasses several points. First the evaluation of the efficacy of the Robo-CSK-Organizer in resolving classification ambiguities within a domestic environment, using its integration of ConceptNet to draw contextual inferences. Secondly, the assessment of the consistency in object placement decisions made by the Robo-CSK-Organizer compared to those made by the ChatGPT Organizer, particularly in how each system handles identical scenarios over repeated trials. Third, the analysis of adaptability in task-relevant classifications, examining how each system adjusts its decision-making when the priority or nature of tasks changes. Finally, a comparison will be drawn on how well both organizers are able to explain their reasoning when sorting.

In regards to delimitations, the research will not extend to other forms of AI systems beyond the Robo-CSK-Organizer and ChatGPT Organizer. The study will not involve human subject testing or surveys, adhering strictly to algorithmic performance assessments.
The environments considered will be simulated domestic settings, and the findings may not directly translate to all real-world domestic environments due to the controlled nature of the experiments.

With these parameters in place, the aim is to provide a focused and in-depth comparison of the two systems' abilities to incorporate commonsense knowledge in a domestic AI application, with an emphasis on ambiguity resolution, consistency, adaptability, and explainability in AI decision-making.

1.8 Definitions of Terms

Commonsense Knowledge (CSK): The collection of facts and information that an average person is expected to know.

Ambiguity Resolution: The process of determining the most appropriate classification for an object that could belong to multiple categories.

Consistency: The ability of a system to provide the same output when presented with the same context multiple times.

Task Relevance: The capacity of a system to alter its classification decisions based on the current priority or task at hand.

Explainable AI (XAI): The field of artificial intelligence focused on designing AI systems whose actions can be easily understood and interpreted by humans.

ConceptNet: A semantic network designed to aid computers in understanding the meanings of words and phrases humans use.

Deep Learning: A subset of machine learning in artificial intelligence that uses algorithms inspired by the structure and function of the brain's neural networks.

Natural Language Processing (NLP): The technology used to aid computers in understanding, interpreting, and manipulating human language.

Human-Robot Interaction (HRI): The study of interactions between humans and robots, focusing on the design, understanding, and evaluation of robotic systems for use by or with humans.

1.9 Thesis Structure

The thesis is organized into five main chapters, supplemented by preliminary sections and appendices that provide additional context and support for the research.

- **Preliminaries**: The thesis begins with an Abstract that succinctly summarizes the study, followed by optional sections for Dedication and Acknowledgments. The Table of Contents, List of Figures, and Acronyms provide an overview and easy navigation of the thesis content.
- **Chapter 1:** Introduction: This chapter sets the stage for the research, outlining the background, problem statement, objectives, significance, contributions, research questions, scope, definitions of terms, and the structure of the thesis.
- **Chapter 2:** Literature Review: An extensive review of existing literature covers the foundational concepts of CSK in AI, human-robot collaboration, advances in domestic robotics, image classification techniques, and a comparative analysis of existing solutions, culminating in a gap analysis that this thesis addresses.

- **Chapter 3:** Methodology: The research methodology is detailed here, including the proposed approach, frameworks for the Robo-CSK-Organizer and ChatGPT Organizer, comparative metrics, implementation strategies, data collection and analysis methods, validation of methods, and ethical considerations.
- **Chapter 4:** Experimental Evaluation: This chapter presents the experimental setup and the results of experiments designed to test ambiguity resolution, consistency, task-relevance adaptability, and explainability. It includes a comparative analysis and discussion of results.
- **Chapter 5:** Conclusions and Future Work: The final chapter summarizes the findings, discusses the limitations of the study, offers recommendations for future research, outlines the implications of the research, and reflects on the research process.
- **Supplementary Sections:** The thesis concludes with a comprehensive list of Works Cited, Appendices containing code listings, raw data and appendix sections such as a Glossary and Index. Additional considerations for ethical implications, quality of visual elements, and formatting consistency are also included to ensure the thesis meets academic standards.

Chapter 2: Literature Review

2.1 Commonsense Knowledge in AI through Knowledge graphs

2.1.1 The Emergence of Commonsense Knowledge in AI

Commonsense knowledge, inherently subtle and nuanced in human cognition, presents unique challenges and opportunities for AI. Unlike encyclopedic knowledge, which is factual and explicit, CSK involves understanding and processing general knowledge about the world that humans consider obvious or "common sense." This distinction is critical in AI's journey, as it highlights the shift from merely processing factual data to interpreting and reasoning with information in a way that aligns with human understanding.

2.1.2 Overview and Historical Development of Knowledge Graphs

Commonsense knowledge (CSK) in AI, especially in the context of knowledge graphs (KGs), involves representing and processing everyday facts and assumptions about the world that humans typically take for granted. KGs have evolved from early semantic networks, which were directed labeled graphs used in AI to represent knowledge. This evolution includes the adoption of first-order logic and relational structures to represent complex relationships and taxonomies. Early KGs were often manually curated and focused on capturing complex axioms for modeling human reasoning, but with the advent of the internet, there was a shift towards capturing vast amounts of ground facts for applications like search and analytics. This shift was less about complex inference and more about organizing and integrating large datasets[5][36].

2.1.3 Key Methodologies and Technologies

KGs are directed labeled graphs where nodes represent entities such as people, companies, or concepts, and edges represent the relationships between these entities.

Technologies like the Resource Description Framework (RDF) and property graph models are used to organize and represent this information. These graphs can be constructed through a combination of human-driven, semi-automated, or fully automated methods, with information added to them being verifiable and easily understood. Modern KGs are not only used for representing structured knowledge but also play a crucial role in AI systems, serving as input and output for machine learning algorithms **See Figure 1**. Technologies like ConceptNet and DBpedia represent vast repositories of commonsense knowledge and semantic networks, respectively, and play a critical role in the development and application of KGs [5][6].

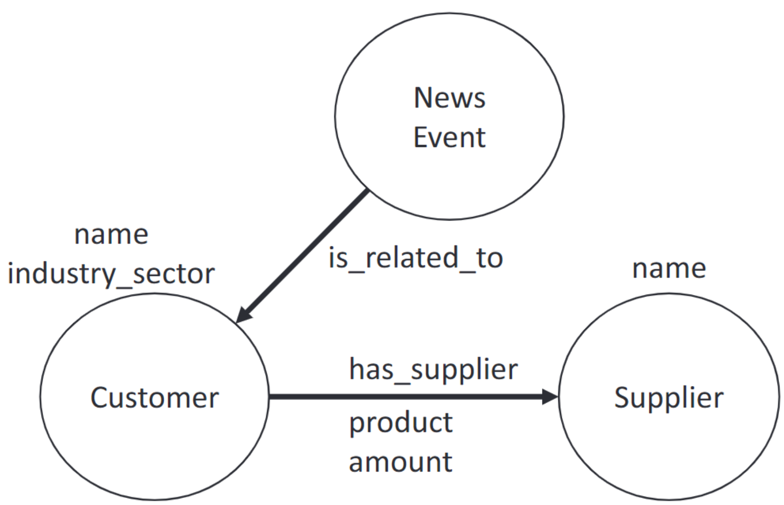


Figure 1: *Example Knowledge Graph* [5]

2.1.4 Applications and Implications

KGs have found widespread applications in various fields. They are crucial in organizing open information on the web, as exemplified by Wikidata, which enhances and improves the quality of information on Wikipedia. In enterprise settings, KGs are used for data integration across various databases and unstructured sources, creating comprehensive views of customer data. They are essential in natural language processing (NLP), computer vision (CV), and commonsense reasoning, where they help in extracting relationships from unstructured text and images. Projects like IBM's Watson have demonstrated the use of KGs in combining symbolic and statistical approaches for commonsense reasoning, an area that continues to evolve with the development of technologies like ATOMIC and GLUCOSE that focus on cause-and-effect reasoning in commonsense knowledge [5].

2.1.5 Challenges and Limitations

The main challenges in developing and maintaining KGs include data sparsity, complexity of graph structures, and ongoing maintenance. The debate between human curation and machine curation highlights the difficulties in balancing automated knowledge extraction with the accuracy and reliability of manually curated data. Additionally, the contrast between "little semantics" (focusing on basic facts) and "big semantics" (capturing deeper meanings of concepts) reflects the inherent limitations in representing complex relationships and concepts within KGs. Resolving these challenges is crucial for advancing KGs' capabilities in AI, particularly in representing commonsense knowledge [5] [44].

KGs have significantly evolved in their application and methodology, contributing immensely to the field of AI, especially in the representation and processing of commonsense knowledge. However, they continue to face challenges that require innovative approaches and solutions.

2.1.6 ConceptNet: A Practical Commonsense Toolkit

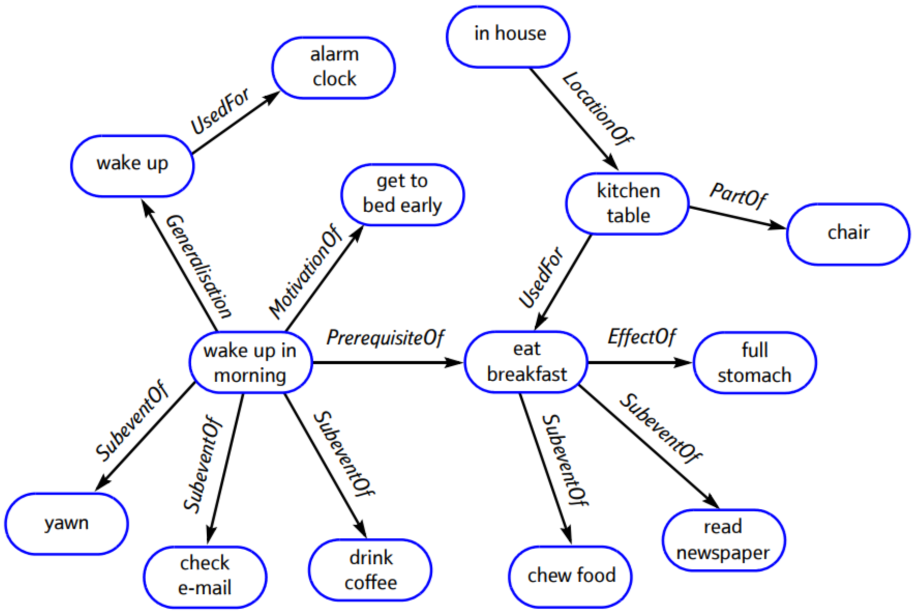


Figure 2: *ConceptNet* [7]

ConceptNet stands out as a significant tool in the realm of KGs **(See Figure 2)**, especially for its application in contextual commonsense reasoning over real-world texts. It is a freely available commonsense knowledge base and natural-language-processing toolkit, supporting practical textual-reasoning tasks over real-world documents. ConceptNet is distinguished by its focus on practical, context-based inferences, leveraging a semantic network of commonsense knowledge. This network currently consists of over 1.6 million assertions of commonsense knowledge, encompassing spatial, physical, social, temporal, and psychological aspects of everyday life [7].

ConceptNet is structurally similar to WordNet but extends beyond lexical items to include higher-order compound concepts composed of action verbs and their arguments, like 'buy food' or 'drive to store.' This expansion allows for a broader range of concepts found in everyday life to be represented. The knowledge in ConceptNet is derived from the Open Mind

Common Sense (OMCS) project, which crowdsourced over 700,000 English sentences of commonsense from the general public. This collaborative approach significantly contrasts with the largely handcrafted nature of resources like WordNet and Cyc. ConceptNet's ontology includes twenty semantic relations, such as EffectOf, SubeventOf, and CapableOf, enhancing its ability to make contextually rich inferences [7][8].

ConceptNet's strength lies in its ability to perform contextual commonsense reasoning, a capability crucial for understanding the gestalt context behind sentences or stories. This ability is crucial for interpreting ambiguous words, understanding sarcasm or irony, and comprehending narratives. ConceptNet's natural language knowledge representation, with nodes as semi-structured English phrases, facilitates this by allowing for the flexibility of meanings and the exploitation of lexical hierarchies. The system's design accommodates the natural ambiguity and redundancy inherent in language, recognizing the value in maintaining different ways of conveying the same idea [7].

In summary, ConceptNet represents a significant advancement in KGs, particularly in the realm of AI and NLP. Its focus on contextual commonsense reasoning, combined with its extensive and collaboratively built knowledge base, makes it a powerful tool for tasks like analogy-making, spatial-temporal-affective projection, and contextual disambiguation. ConceptNet exemplifies the potential for KGs to move beyond structured data representation to more nuanced, context-aware applications in textual information management

2.2 Commonsense Knowledge in AI through Deep Learning and NLP Techniques

2.2.1 Overview and Historical Development:

2.2.1.1 Transition from Traditional Knowledge Bases to AI Systems

As discussed previously, AI systems heavily relied on hand-crafted knowledge bases such as WordNet, ConceptNet and Cyc. Some argue however that these systems, while robust in storing and retrieving factual data, were limited in their ability to process and apply commonsense knowledge in diverse and dynamic scenarios. The realization that CSK involves more than just static facts led to the development of more sophisticated systems capable of dynamic reasoning and learning [9].

2.2.1.2 Deep Learning and NLP: Catalyzing the CSK Evolution

The advent of deep learning and natural language processing (NLP) marked a turning point in the evolution of CSK in AI. These technologies enabled machines to not only store and access vast amounts of data but also to learn and adapt from it, mimicking the human ability to acquire and apply knowledge. This shift from static knowledge bases to dynamic learning systems opened up new possibilities for AI applications, making them more versatile and human-like in their reasoning capabilities [9].

2.2.1.3 Modern AI Systems and CSK

Today, modern AI systems, powered by advanced deep learning models and NLP techniques, have taken CSK to new heights. These systems can process and interpret complex data, understand nuances and context, and even engage in human-like dialogue and decision-making. This progress is not just a testament to technological advancement but also to a deeper understanding of how AI can be designed to think and reason like humans [9].

2.2.2 Key Methodologies and Technologies:

The integration of commonsense knowledge (CSK) into artificial intelligence, specifically through deep learning and natural language processing (NLP) techniques, has witnessed transformative methodologies and technological advancements. This section explores these innovations, focusing on their representation, reasoning, and practical application in AI systems.

2.2.2.1 Commonsense Reasoning in NLP

Commonsense reasoning in NLP has evolved to encompass various methods that enable AI systems to understand and process human-like reasoning. Techniques such as symbolic and statistical approaches have been foundational, but recent advancements in neural network-based methods have significantly improved the ability of machines to mimic human reasoning. These advancements include the development of complex algorithms that process and interpret natural language in a way that mirrors human thought processes [10].

2.2.2.2 Integration of CSK in Deep Learning Models

One notable breakthrough in the application of CSK in AI has been the integration of knowledge bases with deep learning models. The research paper "Augmenting End-to-End Dialogue Systems with Commonsense Knowledge" demonstrates a novel approach to this integration. By incorporating commonsense knowledge bases, such as ConceptNet, into Long Short-Term Memory (LSTM) networks, AI systems can now process and apply CSK more effectively. This integration has been instrumental in enhancing the capability of dialogue systems to understand and respond in a human-like manner [11].

2.2.2.3 Improving Conversational AI with CSK

In enhancing conversational AI, the synthesis of CSK with deep learning models like LSTMs has paved the way for more sophisticated and contextually aware dialogue systems. These systems are not only capable of generating more natural and coherent responses but also demonstrate an improved understanding of the nuances and complexities of human language and thought. The ability to integrate and apply CSK in real-time conversations represents a significant leap in the field of conversational AI [11].

2.2.3 Applications and Implications:

2.2.3.1 Language Understanding and NLP

The integration of commonsense knowledge (CSK) into natural language processing (NLP) has revolutionized language understanding. As outlined in "Commonsense Reasoning for Conversational AI: A Survey," NLP systems now leverage CSK to interpret and generate human-like, contextually relevant responses. This advancement is evident in tasks like sentiment analysis, where understanding the subtleties and nuances of human language is crucial. Moreover, NLP models are increasingly adept at recognizing sarcasm, irony, and metaphor, enriching the interaction between humans and AI [12].

2.2.3.2 Context-Aware Computing

Context-aware computing has been significantly impacted by CSK, particularly in domains like smart home environments and personal digital assistants. CSK enables these systems to interpret human activities, preferences, and intentions more accurately, leading to more intuitive and efficient user experiences. For instance, in smart home environments, systems can make decisions on energy conservation or user comfort by understanding typical human behavior patterns in different contexts, as discussed in "Large Language Models as Commonsense Knowledge" [1]).

2.2.3.3 Conversational AI and Dialogue Systems

In conversational AI, the use of CSK in dialogue systems has led to more natural and coherent conversations. As highlighted in "Commonsense Reasoning for Conversational AI: A Survey," these systems can now engage in complex interactions, including managing ambiguous queries and providing contextually appropriate responses. This improvement enhances user satisfaction and broadens the scope of applications for conversational agents in customer service, therapy, education, and more [12].

2.2.3.4 Broader Implications

The implications of these advancements extend beyond specific applications. By infusing AI systems with CSK, there is a fundamental shift towards creating AI that can operate more autonomously and make decisions that align closely with human reasoning and values. This shift is critical for the future of AI in complex, real-world scenarios where nuanced understanding and decision-making are paramount.

The applications and implications of CSK in AI, particularly in NLP, context-aware computing, and conversational AI, are vast and transformative. These advancements not only enhance the capabilities of individual applications but also mark a significant step toward creating more intelligent, empathetic, and human-aligned AI systems.

2.2.4 Challenges and Limitations:

2.2.4.1 Black-Box Nature of Algorithms

The black-box nature of deep learning and NLP algorithms remains a significant challenge. As "Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction" points out, the decision-making processes of these sophisticated models, especially deep neural networks, are often opaque. This lack of transparency is a major issue in fields that demand explainability, like healthcare and finance, where understanding the rationale behind AI decisions is crucial [13].

2.2.4.2 Data Bias and Ethical Considerations

The issue of data bias in machine learning algorithms is highlighted in "Recent Advances in Natural Language Inference: A Survey." These biases in the training data can lead to skewed or unethical AI outcomes, especially in human-centric applications where such biases can have profound social implications [14].

2.2.4.3 Inexperience in Commonsense Reasoning

The paper "ChatGPT is a Knowledgeable but Inexperienced Solver" sheds light on a specific limitation of large language models (LLMs) like ChatGPT in handling commonsense reasoning. Despite being knowledgeable, these models struggle to identify and apply the necessary commonsense knowledge for specific questions. This limitation highlights the challenges LLMs face in effectively leveraging commonsense knowledge in context, a key aspect of AI applications requiring nuanced understanding[15].

2.2.4.4 Need for Large Datasets

Both the papers discussed above emphasize the reliance on extensive datasets for training AI models. The effectiveness of deep learning and NLP systems depends heavily on the volume and quality of the training data. However, gathering such large datasets can be expensive and challenging, compounded by privacy and data protection concerns in sensitive areas [13][14].

The advancements in AI through deep learning and NLP techniques have been remarkable but come with significant challenges. These include the black-box nature of algorithms, data bias, the need for large datasets, and the inexperienced handling of commonsense reasoning by advanced models like ChatGPT. Addressing these challenges is essential for the ethical, fair, and transparent application of AI technologies across various sectors.

2.2.5 ChatGPT: A Pioneering Tool in AI-Driven Language Processing

ChatGPT, developed by OpenAI, has emerged as a groundbreaking artificial intelligence (AI) language model, reshaping the landscape of scientific research and application domains such as customer service, healthcare, and education. This model is part of the generative AI class, which creates new data based on existing patterns and structures, relying on deep learning and neural networks. Distinguished by its ability to understand and generate human-like text, ChatGPT represents a significant leap in natural language processing (NLP) capabilities [16].

The inception of ChatGPT can be traced back to the development of the Transformer architecture, a fundamental shift in NLP that overcame limitations of previous models like RNNs and CNNs. Built on the GPT-3.5 architecture, a modification of the original GPT-3, ChatGPT, although smaller in parameters, excels in tasks such as language understanding, text generation, and machine translation. Its training involved a large corpus of text data, specifically fine-tuned for generating conversational responses, thereby enhancing its human-like interaction quality [16].

2.2.5.1 Key Innovations and Strengths:

Enhanced Contextual Understanding: ChatGPT's ability to comprehend complex and nuanced inputs is a testament to its advanced understanding of context, essential for generating accurate and relevant responses [16]).

Bias Reduction and Fine-Tuning: Continuous efforts are made to minimize biases in its training data, contributing to more balanced outputs. The model's fine-tuning capabilities allow it to be tailored to specific tasks and applications, catering to diverse scientific needs [16].

Resolving Conversational AI Challenges: ChatGPT addresses several conversational AI challenges such as maintaining context over multiple turns, handling ambiguity, personalization, common sense reasoning, and emotional intelligence. These advancements significantly enhance user interaction and experience [16].

Multifaceted Features: The model's versatility is evident in its task adaptability, multilingual proficiency, and scalability. It supports zero-shot and few-shot learning, reducing the need for extensive training and large labeled datasets. These features make ChatGPT highly effective for a range of applications, from customer support to content creation and tutoring [16].

Prompt Engineering: ChatGPT's interaction quality is further improved through prompt engineering, where clear, specific prompts and context-rich information guide the model to generate more accurate and relevant responses [16].

ChatGPT stands as a prominent AI tool, demonstrating exceptional capabilities in language processing and interaction. Its development reflects significant advancements in AI, addressing critical challenges and setting new standards in human-AI collaboration. Despite some controversies and ethical concerns, its impact and rapid adoption in academia, research, and industry sectors are noteworthy, indicating its substantial role in the evolving AI landscape [16].

2.3 Human-Robot Collaboration

2.3.1 Exploration of Human-Robot Interaction (HRI) Principles

Human-robot interaction (HRI) is an evolving field that bridges the gap between humans and robotic technology. The fundamental principles of HRI focus on understanding and enhancing the ways humans and robots can work together effectively. This collaboration is not just a matter of technological interaction but also involves cognitive, social, and emotional aspects.

According to Sheridan's work, HRI encompasses various domains, including supervisory control, teleoperation/telerobotics, automated vehicles, and social interaction. The concept of supervisory control, where humans oversee and intervene in robotic operations, highlights the need for a balance between automation and human decision-making. In teleoperation and telerobotics, humans control robots remotely, often in hazardous or inaccessible environments, emphasizing the importance of intuitive control interfaces and effective communication systems.

Automated vehicles, such as self-driving cars, present challenges in ensuring safety and reliability, requiring robust AI systems that can interact seamlessly with human passengers or operators. Social interaction between humans and robots, particularly in assistive and educational roles, underlines the need for robots to exhibit social behaviors and emotional intelligence [17]. These principles are crucial in guiding the design and deployment of robots in various contexts, ensuring that human-robot collaboration is safe, efficient, and socially acceptable.

2.3.2 Technological Advancements Facilitating HRI

Technological advancements have played a pivotal role in advancing HRI. Key developments include the integration of AI and machine learning, improved sensory and perception systems, enhanced communication interfaces, and advancements in robotics design and engineering.

AI and machine learning algorithms enable robots to interpret and adapt to human behavior, learn from interactions, and make autonomous decisions based on contextual understanding. Sensory and perception systems, such as computer vision and tactile sensors, allow robots to perceive their environment more accurately, facilitating better interaction with humans and the surroundings.

Communication interfaces, including natural language processing and haptic feedback systems, have made interactions with robots more intuitive and human-like. These interfaces allow for more natural communication, enhancing the collaborative experience. The advancements in robotics design, such as the development of soft robotics and bio-inspired designs, have led to more adaptable and safe robots capable of operating in diverse environments and in close proximity to humans [18].

The exploration of HRI principles, coupled with technological advancements, is paving the way for more integrated, effective, and meaningful human-robot collaborations. These developments hold significant promise for various applications, from industrial automation to assistive technologies, ultimately leading to a future where humans and robots work synergistically across different domains [37]. See below for some example case studies of HRI.

2.3.3 Case Studies or Examples of Successful Human-Robot Collaboration

2.3.3.1 Case Study 1: Gesture Recognition for Human-Robot Interaction

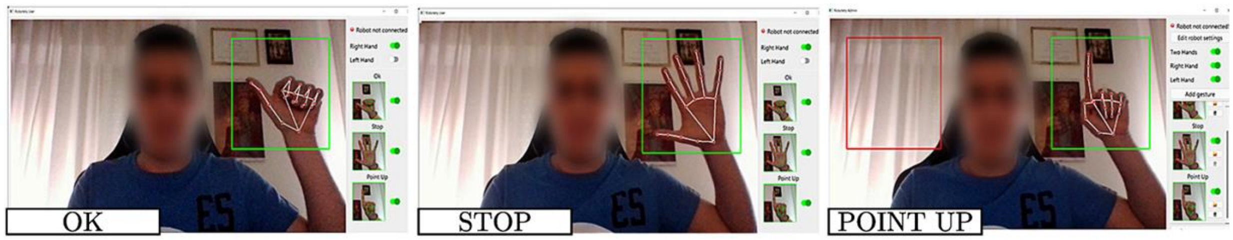


Figure 3: *Gesture Recognition [19]*

A notable example of human-robot collaboration is the use of gesture recognition for interactive communication between humans and robots **(See Figure 3)**. In a study conducted at i-LABS, researchers integrated gesture recognition using a standard webcam and Robotely software, enabling seamless human-robot interaction. This system was effectively implemented with the ABB YuMi cobot (collaborative robot), a programmable system based on RGBD sensors. The technology allowed for the recognition of various hand gestures, which were converted into robot commands like stopping, quality inspection, or manipulating objects like LEGO bricks. This case study demonstrates the potential of simple, intuitive forms of communication in enhancing the efficiency and safety of collaborative tasks in robotic applications[19].

2.3.3.2 Case Study 2: Safe Autonomous Motions and Collision Avoidance Strategies

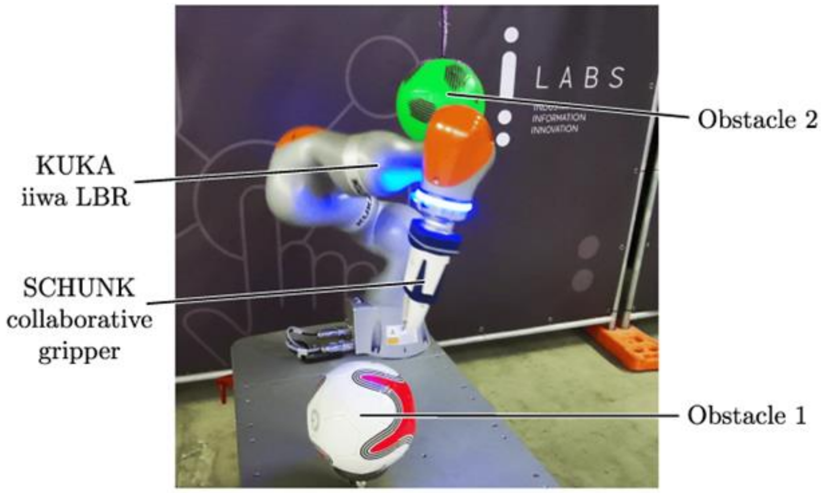


Figure 4: *Collision Avoidance [19]*

Another significant advancement in human-robot collaboration is the development of safe autonomous motions and collision avoidance strategies **(See Figure 4)**. Researchers at i-LABS developed algorithms for obstacle avoidance, which are fundamental for safe human-robot coexistence. These strategies were implemented in a generic manipulator with a high number of actuators and a complex kinematic chain. The algorithms used a Closed-Loop Inverse Kinematic (CLIK) approach and incorporated repulsive velocity fields to prevent collisions. This case study highlights the importance of ensuring safety in collaborative environments, especially where robots and humans share the same workspace. It shows that with intelligent design and robust control algorithms, robots can effectively avoid obstacles while maintaining their task performance [19].

2.3.3.3 Case Study 3: Virtual Reality Based Design Methods for Human-Robot Collaboration

Figure 5: *Virtual Reality HRC [19]*

The integration of Virtual Reality (VR) in designing human-robot collaboration systems is another remarkable case study **(See Figure 5)**. The i-LABS researchers utilized VR as a tool for testing and calibrating control strategies for human-robot interaction. By creating a virtual implementation of the control strategy, they were able to simulate interactions and optimize the robot's responses to dynamic obstacles. This approach serves as an efficient method for designing and testing human-in-the-loop applications, demonstrating the potential of VR tools in enhancing the safety and efficiency of collaborative systems[19].

2.3.4 Ethical and Social Considerations in Human-Robot Interaction (HRI)

2.3.4.1 Ethical Dimensions in HRI

Ethical considerations in HRI are vital for ensuring responsible integration of robotics in society. Ostrowski et al. (2022) stress the importance of ethics in designing and implementing robotic technologies. Their study highlights five ethical senses in HRI: the ethics of roboticists in system design, the programmed moral code of robots, the robot’s autonomous ethical reasoning, ethical relationships between humans and robots, and the broader social concerns around robots. This multi-faceted approach to ethics acknowledges the complexity of HRI, encompassing both the creators’ responsibilities and the societal impact of robotic systems. Addressing these ethical dimensions is crucial for fostering trust and acceptance of robotic technologies in various social contexts[20].

2.3.4.2 Engaging with Societal Problems and Justice in HRI

The societal impact of robots is a significant consideration in HRI. Ostrowski et al. advocate for engaging with societal challenges and justice in the design and deployment of robotic systems. They emphasize the need for HRI research to address social justice issues, ensuring equitable distribution of benefits and burdens of robotic technologies. The study argues for the integration of frameworks like Design Justice in HRI to assess and design for the societal reshaping potential of robotic technologies. This approach is crucial for preventing the reinforcement of existing societal inequities and power imbalances through robotic interventions[20].

2.3.4.3 The HRI Equitable Design Framework

Ostrowski et al. extend the Design Justice framework to HRI, proposing the HRI Equitable Design Framework. This framework guides researchers and practitioners to consider equity and justice in HRI research and practice. It encourages the consideration of various aspects, including who is included in robot design, the intended users, the values and biases embedded in robots, and the societal implications of robotic technologies. By focusing on these aspects, the HRI Equitable Design Framework aims to promote equitable and inclusive HRI studies and technologies, addressing the ethical, social, and justice-oriented concerns in the field [20].

2.3.4.4 Ethical Consideration of Explainable AI in HRI

Explainable AI (XAI) is becoming an integral part of ethical considerations in HRI [45]. Gao et al. (2020) highlight the importance of XAI in facilitating effective communication and collaboration between humans and robots. Their approach involves robots building hierarchical mind models of human users and generating explanations of their own mental states, which enhances collaboration performance and user perception of the robot. This form of goal-driven XAI increases model transparency, fosters human trust, and improves task performance, thereby addressing ethical concerns about AI being a "black box." In practice, these explanations help align human and robot mental states, correct misunderstandings, and provide users with necessary task-related knowledge, making robot behavior more understandable and acceptable to human collaborators[21].

2.4 Advances in Domestic Robotics

2.4.1 Historical Progression and Current State of Domestic Robotics

The historical progression and current state of domestic robotics can be traced back to its origins as a concept in science fiction and early conceptualizations in plays like Karel Capek's "Rossum's Universal Robots." The idea of a robot performing household chores has been a long-standing aspiration, but the reality has been more gradual and complex [22].

Initially, domestic robots were envisioned as multifunctional devices capable of performing a wide range of household tasks, such as cleaning, cooking, and other chores. However, despite the significant potential customer base and market opportunity, the development and commercialization of such robots have been slow and fraught with challenges, both technical and economic [22].

The current state of domestic robotics includes advancements in specific areas like robotic lawn mowing, ironing robots, intelligent refrigerators, and digital wardrobes. These developments indicate a shift from the idea of a single, do-it-all domestic robot to specialized appliances that focus on particular tasks [22].

Economically, the professional cleaning market, which includes services in Europe alone, was worth approximately $50 billion annually as of 1995. A significant portion of this market could potentially be automated using robotic solutions. The domestic market also presents a significant opportunity, as evidenced by the success of products like the Roomba, which brought robotic vacuum cleaners into mainstream awareness after its launch in 2002 [22].

2.4.2 Technological Innovations Driving Advancements in Domestic Robots

The realm of domestic robotics has witnessed significant technological innovations that have been pivotal in shaping the current landscape of this field. These advancements have expanded the capabilities of domestic robots, enabling them to perform a wide array of household tasks with greater efficiency and autonomy.

One of the key technological advancements is in the area of sensor technology. Modern domestic robots are equipped with advanced sensors that allow for precise navigation, obstacle detection, and environment mapping. This technology is crucial for tasks such as vacuuming, where the robot must navigate around furniture and other obstacles while ensuring thorough coverage of the cleaning area [22].

Another significant innovation is in the development of machine learning algorithms. These algorithms enable domestic robots to learn from their environment and adapt their behavior over time. For example, a robotic vacuum cleaner can learn the layout of a house and optimize its cleaning path accordingly. This adaptability not only improves the efficiency of the robot but also enhances its usability in diverse household settings [22].

Moreover, advancements in battery technology and energy efficiency have been crucial. Modern domestic robots are now capable of longer operation times and require less frequent charging. This improvement in energy efficiency means that robots can perform tasks such as lawn mowing or floor cleaning for extended periods, thereby increasing their practicality for everyday use [22].

In addition to these technological advancements, there has also been significant progress in the area of human-robot interaction. Contemporary domestic robots are often equipped with user-friendly interfaces, making them accessible to a broader range of users, including those without technical expertise. These interfaces may include voice control, smartphone integration, and intuitive touch controls, making the interaction with the robot more natural and efficient [22].

Lastly, the integration of robotics with smart home technologies has opened up new possibilities for domestic robots. By connecting with other smart devices in the home, these robots can operate more synergistically within the household ecosystem. For example, a robot can start cleaning when it detects that the house is empty or can communicate with other smart appliances to optimize household chores [22].

The advances in sensor technology, machine learning, battery life, human-robot interaction, and integration with smart home systems have collectively driven the evolution of domestic robotics. These technological innovations have not only enhanced the capabilities and efficiency of domestic robots but have also played a key role in their increasing adoption in households worldwide [39].

2.4.3 Impact of Domestic Robots on Everyday Life

The impact of domestic robots on everyday life has been profound and multifaceted, transforming how household tasks are approached and contributing significantly to the evolution of smart home technologies. This section explores the influence of these robots through various case studies and examples, demonstrating their role in modern households.

One notable impact is the automation of routine household chores, which has been a major relief for many homeowners. The introduction of robotic vacuum cleaners, such as the Roomba, exemplifies this shift. These robots have automated the task of floor cleaning, a chore that was traditionally time-consuming and labor-intensive. With the ability to navigate around obstacles and clean under furniture, robotic vacuum cleaners have made floor cleaning more efficient and less burdensome for homeowners [22].

Another area where domestic robots have made a significant impact is in lawn maintenance. Robotic lawn mowers have revolutionized this task by automating the process, enabling lawns to be maintained with minimal human intervention. These robots operate autonomously, adjusting their mowing patterns to the size and shape of the lawn, and can return to their charging stations when their battery runs low. This advancement not only saves time but also contributes to a more consistent and aesthetically pleasing lawn appearance [22].

Moreover, domestic robots have also ventured into more specialized tasks, such as ironing and clothes management. The development of ironing robots and digital wardrobes has begun to transform how households manage their clothing. These technologies offer the potential to automate the process of ironing, folding, and organizing clothes, reducing the time and effort typically associated with these tasks [22].

In addition to task-specific impacts, domestic robots have also contributed to the overall advancement of smart home ecosystems. By integrating with other smart home devices, these robots can operate more intelligently and efficiently. For instance, a robotic vacuum cleaner can be programmed to start cleaning when the smart home system detects that the house is empty, optimizing energy usage and minimizing disruption to the inhabitants [22].

Finally, the psychological and social implications of domestic robots cannot be overlooked. These robots not only alleviate physical labor but also contribute to the emotional well-being of homeowners by reducing the stress associated with household chores. The sense of comfort and convenience provided by these robots has led to a more positive perception of household management and an increased sense of control over one's living environment [22].

The impact of domestic robots on everyday life is evident across various household tasks. From automating routine chores like cleaning and lawn maintenance to integrating with smart home systems, these robots have not only enhanced the efficiency of household management but also contributed to the well-being of homeowners. As technology continues to evolve, it is likely that the influence of domestic robots will only deepen, further transforming the landscape of modern living.

2.4.4 Future Trends and Potential Developments in Domestic Robotics

Figure 6: *Samples of Domestic Robots [23]*

The future of domestic robotics is shaped by several emerging trends and potential developments, each contributing to the increasing sophistication and utility of these technologies in everyday life **(See Figure 6)**.

2.4.4.1 Integration of Advanced Computational Intelligence and Multi-Modal Learning:

Future domestic robots are expected to incorporate advanced computational intelligence models, including multi-modal learning and meta-cognition. These technologies will enable robots to process and understand complex information from various sensory inputs more effectively, leading to more intuitive and human-like interactions [23].

2.4.4.2 Expansion of Market and Range of Functions:

With the rapid advancement in machine learning, faster internet connections, and more affordable hardware, the market for domestic robots is projected to grow significantly. Future robots will likely be designed to perform a variety of physical tasks, communicate more effectively with users, and provide social companionship, effectively blending into the domestic environment [23].

2.4.4.3 Bridging the Gap Between Academic Models and Commercial Products:

There is a current gap between the most advanced computational intelligence models developed in academia and the technologies implemented in commercial domestic robots. Bridging this gap is crucial for the next generation of domestic robots, which will require balancing the cutting-edge capabilities with practicality and user safety [23].

2.4.4.4 Diversification of Robot Types and Functions:

The future will likely see a broader spectrum of domestic robots, ranging from virtual assistants and IoT robots to interactive and service robots. Each category will serve specific roles, from managing smart home devices to providing physical assistance and social interaction [23].

2.4.4.5 Development of Multipurpose Domestic Robots:

The concept of multipurpose domestic robots, which combine various functionalities and seamlessly integrate with other smart home devices, is gaining traction. These robots will be capable of performing diverse tasks, from controlling home appliances to providing cognitive assistance and social interaction [23].

2.4.4.6 Improvement in Core Functions of Robots:

Future domestic robots are expected to improve significantly in core functions such as perception, action, understanding, and communication. This will encompass advanced SLAM (Simultaneous Localization and Mapping) for navigation, more effective obstacle avoidance, enhanced action and emotion recognition, and more sophisticated communication abilities including natural language processing and understanding [23]).

2.4.4.7 Focus on Data Safety and Ethics:

As domestic robots become more integrated into daily life, issues of data safety, ethics, and explainability will become increasingly important. Ensuring that these robots operate within ethical boundaries and maintain user privacy will be a critical aspect of their development and acceptance [23].

The future of domestic robotics is marked by a blend of technological advancements, market expansion, and an increased focus on user interaction, data safety, and ethical considerations. These developments promise to make domestic robots an even more integral and versatile component of everyday life.

2.5 Advanced Image Classification and Recognition Technologies

2.5.1 Detailed Examination of Key Technologies in Image Classification and Recognition

In the rapidly progressing field of image classification and recognition, deep learning has emerged as a cornerstone technology, revolutionizing the sector with its accuracy and efficiency. Currently the pivotal technologies at the forefront of image recognition are Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs). Each technology brings distinct advantages, from processing complex image data to managing sequential data and innovatively generating new samples. These technologies have found significant applications in essential areas such as face recognition, medical imaging, and remote sensing, highlighting their adaptability and effectiveness. Yet, challenges like the requirement for extensive training data, high computational demand, and the need for integration into smaller devices remain, directing the future path of research and innovation in this domain.

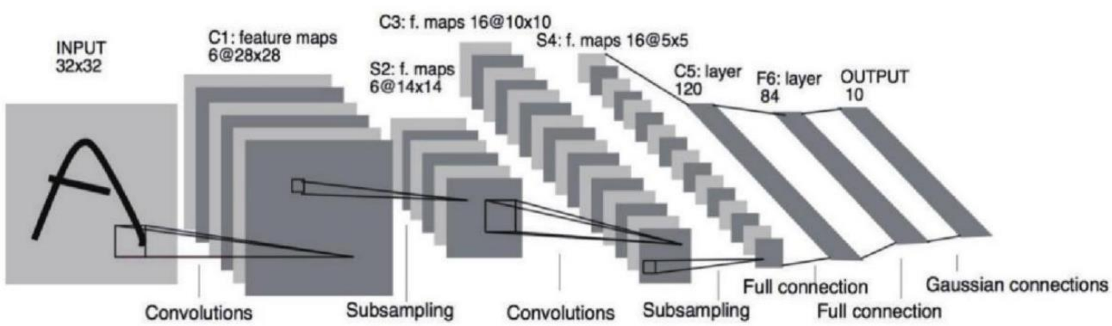


Figure 7: Convolutional Neural Network (CNN) [26]

Convolutional Neural Networks (CNNs): CNNs have revolutionized the field of image recognition by streamlining the image pre-processing step and reducing the need for manual feature extraction. They operate on the assumption of local correlation within images, focusing on neighboring pixels' correlation. This local connectivity feature enables CNNs to efficiently process and classify images by automatically extracting and learning features from raw image data **(See Figure 7)** [24].

Recurrent Neural Networks (RNNs): RNNs are designed to handle sequential or time-series data. They are characterized by their memory function, where the output of a neuron includes not just the input from the previous layer but also the output from its prior input. This makes RNNs particularly suitable for applications where data points are interdependent, although they are less common in static image recognition tasks [24].

Generative Adversarial Networks (GANs): GANs consist of two models, a generative model and a discriminative model, that work in tandem through an adversarial process. The generative model creates new data samples, while the discriminative model evaluates them against real data. This innovative approach allows for more efficient and varied data generation and has found widespread application in image processing [24][35].

2.5.1.1 Applications in Specific Fields:

Face Recognition: Deep learning has significantly enhanced face recognition technology, allowing for more accurate and efficient identification based on human facial characteristics. Algorithms based on deep learning have shown greater accuracy in face recognition, even under challenging conditions like occlusions and variations in expression and lighting [24].

Medical Image Recognition: In the medical field, CNNs are the preferred algorithm for image recognition, enabling accurate diagnosis and treatment planning. They excel in extracting and processing pathological information from medical images, such as identifying lesions in various body parts, including the brain, chest, and abdomen [24].

Remote Sensing Image Recognition: Deep learning techniques effectively process remote sensing images by automatically extracting spatial and semantic information. They have improved the accuracy of identifying features like objects, buildings, roads, and natural resources and are particularly effective in reducing data dimensionality in large remote sensing datasets [24].

2.5.1.2 Future Development and Challenges:

Despite its successes, deep learning in image recognition still faces challenges. These include the need for large training datasets, the high computational complexity of models, and the necessity of deploying these models on lightweight devices. Future research directions include optimizing training data, reducing model complexity, and exploring unsupervised and self-supervised learning methods to achieve satisfactory results under limited conditions [24].

2.5.2 Comparative Analysis of Different Technologies Used in Image Classification

The field of image classification and recognition has been profoundly transformed by the advent of deep learning technologies. Among these, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs) stand out as key players, each offering unique capabilities and strengths. CNNs, with their multi-layered architecture, are adept at processing and interpreting image data. RNNs, on the other hand, excel in handling sequential and time-series data, offering an internal memory feature that captures context over time. GANs, a more recent innovation, revolutionize data generation and assessment with their dual-model structure. Together, these technologies represent the diverse and dynamic nature of deep learning in the realm of image recognition, highlighting its versatility and potential for continual advancement.

Convolutional Neural Networks (CNNs): CNNs are pivotal in the realm of image recognition, characterized by their multi-layered network structure mimicking the human brain's neural setup. The strength of CNNs lies in their layered architecture, typically comprising an input layer, followed by alternating convolutional and pooling layers, and concluding with a fully connected multilayer perceptron classifier. This structure is adept at processing image data, allowing CNNs to efficiently identify patterns and features in images. The evolution of CNNs over time has led to significant improvements in both the architecture and functionality of these networks, making them a fundamental tool in image recognition [25].

Recurrent Neural Networks (RNNs): In contrast to CNNs, RNNs are designed to handle sequential data, making them less common in static image recognition tasks but invaluable in applications involving time series or sequential data. RNNs have a unique 'memory' feature where the output is influenced by both the current input and the network's previous state. This characteristic allows RNNs to maintain a form of internal state that captures information about previous inputs, making them suitable for tasks where context or order is important [26].

Generative Adversarial Networks (GANs): GANs are a more recent development in the field of deep learning, known for their two-model structure consisting of a generative and a discriminative model. This unique framework allows GANs to generate new data samples and simultaneously evaluate them, a feature particularly useful in image generation and enhancement tasks. GANs have brought a new dimension to image processing with their ability to create and assess data, demonstrating deep learning's expanding capabilities [24].

Each of these technologies offers distinct advantages in the field of image classification and recognition. While CNNs are lauded for their direct applicability to image data, RNNs excel in sequential data analysis, and GANs introduce innovative data generation and evaluation capabilities. This diversity in technologies underlines the versatility and depth of deep learning applications in image recognition.

2.5.3 Integration and Application of Detic and BLIP in Object Detection and Environmental Context Understanding

Detic's Role in Object Detection: Detic is a novel object detector that significantly expands the vocabulary of detectors to tens of thousands of concepts, addressing the limitations of traditional object detectors. Current detectors are constrained by the smaller scale of detection datasets, whereas image classifiers manage larger vocabularies due to their larger and easier-to-collect datasets. Detic integrates these two by training the classifiers of a detector on image classification data. This approach simplifies the process by eliminating the need for complex assignment schemes to assign image labels to boxes, making it easier to implement and compatible with various detection architectures. Detic excels in detecting classes without box annotations and shows superior performance on both open-vocabulary and long-tail detection benchmarks, closing the performance gap for object categories with few samples [27].

BLIP's Contribution to Environmental Context Understanding: BLIP (Bootstrapping Language-Image Pre-training) is a Vision-Language Pre-training (VLP) framework designed to enhance the performance of many vision-language tasks. It addresses the limitations of existing models that are either focused on understanding-based or generation-based tasks. BLIP effectively utilizes noisy web data by generating synthetic captions and filtering out the noisy ones. This approach has led to state-of-the-art results in a variety of vision-language tasks like image-text retrieval, image captioning, and visual question answering (VQA). BLIP's strong generalization ability also extends to video-language tasks in a zero-shot manner. This framework employs a visual transformer as the image encoder, a multimodal mixture of encoderdecoder (MED) model for multi-task pre-training, and optimizes three objectives: Image-Text Contrastive Loss (ITC), Image-Text Matching Loss (ITM), and Language Modeling Loss (LM). The use of CapFilt, a method for improving the quality of the text corpus by filtering out noisy texts, further enhances BLIP's performance [28][46].

Synergistic Integration of Detic and BLIP: In the context of robotic applications, integrating Detic and BLIP presents a comprehensive solution for object detection and environmental context understanding. Detic's advanced object detection capabilities, combined with BLIP's proficiency in interpreting environmental contexts through language-image pre-training, provide robots with a nuanced understanding of their surroundings. This integration enables robots not only to detect and localize objects with high accuracy but also to understand the nature of the environment they are operating in, whether it's a living room or a bathroom. The combined strengths of Detic in detailed object identification and BLIP in contextual understanding pave the way for more intelligent and capable robotic systems, enhancing their ability to interact with and adapt to diverse domestic environments.

2.6 Comparative Analysis of Existing Solutions for CSK integration in Robotics

The following segment delves into a detailed examination of various methodologies and approaches in the realm of commonsense knowledge (CSK) integration within robotic systems, particularly in domestic environments. This analysis reviews existing solutions, providing a comprehensive understanding of the current state of CSK in robotics. Additionally, it discusses diverse approaches to similar problems, highlighting the effectiveness and innovation in CSK application. This exploration is pivotal in contextualizing the contributions of this thesis within the broader landscape of robotics research, showcasing how the proposed solutions compare and potentially advance beyond current methodologies. The insights gleaned from this comparative study are crucial for understanding the complexities and challenges in integrating CSK into robotic systems. They set the stage for the introduction of the innovative Robo-CSK-Organizer system, representing a significant advancement in the field.

2.6.1 Review of Existing Solutions

2.6.1.1 Commonsense Knowledge Extraction for Tidy-up Robotic Service in Domestic Environments:

This paper presents a novel approach to integrating commonsense knowledge into domestic robotics, focusing on object classification and handling for a tidy-up service. The solution employs ConceptNet and Google search data to categorize objects into functional groups, enhancing the robot's understanding and interaction with its environment [29].

2.6.1.2 Large Language Models as Commonsense Knowledge:

The study explores the utilization of large language models, specifically GPT-3.5, as a repository of commonsense knowledge for task planning in robotics. This approach demonstrates the potential of language models to enrich robotic decision-making processes with a broad base of CSK, thereby improving their adaptability and functionality in domestic settings [1].

2.6.1.3 Putting People's Common Sense into Knowledge Bases of Household Robots:

This paper discusses the incorporation of commonsense knowledge from the OMICS database into household robots, utilizing Description Logic for knowledge representation. This integration allows robots to perform more context-aware and nuanced tasks within a domestic environment [30].

2.6.1.4 Robot Action Planning by Commonsense Knowledge in Human-Robot Collaborative Tasks:

The research focuses on the application of CSK in human-robot collaborative tasks, particularly in robot action planning for assembly tasks. It highlights how CSK can enhance cooperative interactions and decision-making in complex task scenarios [31].

2.6.1.5 Semantic Task Planning for Service Robots in Open World:

This study presents a system for semantic task planning in service robots, leveraging natural language understanding and semantic reasoning. It addresses the challenges of dynamic, open-world environments, showcasing how CSK can enhance a robot's ability to adapt and respond to changing situations [2].

2.6.1.6 Commonsense Reasoning and Knowledge Acquisition to Guide Deep Learning on Robots:

This paper investigates the integration of non-monotonic logical reasoning and incomplete commonsense domain knowledge with inductive learning to guide deep learning on robots. It offers a unique perspective on the confluence of CSK and advanced learning techniques for robotic tasks [3].

2.6.2 Discussions on Approaches to CSK Integration In Robotics

2.6.2.1 Approach from 'Commonsense Knowledge Extraction for Tidy-up Robotic Service in Domestic Environments':

This study leverages web data and ConceptNet to build CSK for domestic robotics, particularly for object classification and handling. The approach is unique in its combination of web data with structured commonsense knowledge, providing a robust foundation for decision-making in robots. This methodology is effective for categorizing objects into functional groups, crucial for task-specific operations like tidying up [29].

2.6.2.2 Use of Large Language Models in 'Large Language Models as Commonsense Knowledge':

In this research, large language models (LLMs) such as GPT-3.5 are utilized for CSK extraction and task planning. LLMs offer a vast repository of knowledge and reasoning capabilities, making them highly effective in processing complex tasks and providing contextual understanding. This approach showcases the potential of LLMs in enhancing the decision-making processes of robots, particularly in task planning and execution [1].

2.6.2.3 CSK Integration in 'Putting People’s Common Sense into Knowledge Bases of Household Robots':

This paper discusses the integration of CSK from the OMICS database into household robots using Description Logic. This method stands out for its ability to transform natural language into formal knowledge representations, enabling robots to perform nuanced tasks with a deeper understanding of their environment [30]).

2.6.2.4 Human-Robot Collaboration in 'Robot Action Planning by Commonsense Knowledge in Human-Robot Collaborative Tasks':

The research focuses on CSK application in human-robot collaborative tasks. By employing commonsense knowledge in action planning, robots can better understand and anticipate human actions, leading to more effective collaboration in tasks such as assembly [31].

2.6.2.5 Semantic Task Planning from 'Semantic Task Planning for Service Robots in Open World':

This study presents a system for semantic task planning in service robots, leveraging natural language understanding and semantic reasoning. It addresses the challenges of open-world environments, highlighting how CSK can enhance a robot's adaptability and responsiveness [2].

2.6.2.6 Deep Learning and CSK in 'Commonsense Reasoning and Knowledge Acquisition to Guide Deep Learning on Robots':

This paper investigates combining non-monotonic logical reasoning with CSK for guiding deep learning in robotics. The integration of reasoning and learning methods provides a unique perspective on using CSK for complex decision-making, particularly in learning and adapting to new environments and tasks [3].

Each of these approaches contributes to the field of domestic robotics by offering innovative solutions to integrate CSK. They demonstrate effectiveness in areas such as task planning, human-robot interaction, and environmental adaptation. These methodologies provide valuable insights into the diverse ways CSK can be utilized to enhance the functionality and intelligence of robotic systems in domestic settings.

2.6.3 Comparison with Robo-CSK-Organizer

2.6.3.1 Integration of Detic and BLIP in Robo-CSK-Organizer

Robo-CSK-Organizer not only integrates ConceptNet but also employs Detic and BLIP, creating a synergistic framework that significantly enhances object detection and environmental context understanding. Detic expands the object detection capabilities of Robo-CSK-Organizer by incorporating a vast vocabulary of concepts. Its ability to detect a wide array of objects, including those with few samples, addresses the limitation of traditional object detectors. This aspect of the Robo-CSK-Organizer stands out compared to existing solutions that may struggle with the detection of less common or ambiguous objects in domestic environments [27].

BLIP, with its Vision-Language Pre-training, augments the Robo-CSK-Organizer's environmental understanding. It goes beyond mere object recognition, allowing the system to interpret and understand the context of environments, such as distinguishing between a kitchen and a bedroom. This ability to process and integrate visual and linguistic information sets the Robo-CSK-Organizer apart from others that may lack nuanced environmental context understanding [28].

2.6.3.2 Distinctive Features of Robo-CSK-Organizer

Comprehensive Object Classification and Placement: Combining Detic, BLIP, and ConceptNet, the Robo-CSK-Organizer can classify and sort a vast range of objects accurately. Its utilization of commonsense knowledge from ConceptNet, enhanced by the detailed object identification from Detic and contextual understanding from BLIP, enables the system to place objects in the most contextually appropriate locations.

Advancements in Ambiguity Resolution: The integration of these advanced technologies allows the Robo-CSK-Organizer to effectively resolve ambiguities in object placement, a challenge that has been a significant limitation in previous research. Where other systems may falter in ambiguous or complex scenarios, the organizer manages to excel at these.

Enhanced Real-World Adaptability and Explainability: The combined use of these advanced tools improves the system's adaptability to real-world environments. The inclusion of ConceptNet contributes to the system’s ability to explain its decisions and actions, a feature that is often lacking in other systems. This aspect is crucial for user interaction and trust.

2.6.3.3 Comparison Over Existing Research

In comparison to existing solutions, the Robo-CSK-Organizer's approach to object classification and environmental understanding is more holistic and nuanced. While previous research has made significant strides in either object detection or CSK integration, the Robo-CSK-Organizer system marries these aspects with advanced language and image processing capabilities. This integration results in a system that not only understands what objects are but also comprehends where they belong in the context of a dynamic, real-world environment, addressing a critical gap in existing domestic robotics research.

The Robo-CSK-Organizer represents a significant leap forward in the realm of domestic robotics, offering a comprehensive, intelligent, and adaptable solution for object organization and classification. Its unique integration of Detic, BLIP, and ConceptNet places it at the forefront of current technological advancements, distinctly differentiating it from the existing body of research.

2.7 Theoretical Framework

This thesis integrates a range of theories and models from artificial intelligence (AI), robotics, and commonsense knowledge. Notably, it leverages the principles of machine learning, particularly deep learning, as foundational to AI advancements. Models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are central to the Robo-CSK-Organizer's design, enabling sophisticated object recognition and classification. Additionally, the study draws upon the concepts of human-robot interaction (HRI), emphasizing the importance of intuitive, context-aware AI systems in domestic settings.

The integration of commonsense knowledge (CSK) in robotics, central to this research, is underpinned by the synthesis of AI and human-like reasoning models. The Robo-CSK-Organizer, developed as part of this study, embodies the application of these theories. It utilizes ConceptNet, a semantic network, to imbue robots with human-like contextual understanding, directly addressing the research objectives of enhancing decision-making and interaction capabilities in multipurpose robots.

The theoretical underpinnings of this research have shaped its methodological approach. The comparative analysis framework and the criteria for evaluating the experiments are derived from principles of XAI (Explainable AI), emphasizing the need for transparency and interpretability in AI systems. The development of the Robo-CSK-Organizer is guided by the principle that AI should be both effective in task performance and comprehensible to users.

This research contributes to existing theories by demonstrating the practical application and benefits of integrating CSK in AI, particularly in robotics. It offers insights into the application of commonsense reasoning in AI, challenging the conventional approach of relying solely on algorithmic decision-making. This study provides empirical evidence of the feasibility and efficiency of integrating CSK in AI, potentially influencing future theoretical developments in AI and robotics.

The findings of this research have significant implications for the future of AI and robotics theory. They suggest new avenues for integrating CSK in various AI applications, potentially leading to more intuitive and human-centered AI systems. This research could spur further theoretical advancements in the field, promoting the development of AI systems that are not only intelligent but also possess a depth of understanding akin to human commonsense reasoning.

2.8 Summary of the Literature Review

The literature review in this thesis encompasses a broad spectrum of research in the fields of artificial intelligence (AI), commonsense knowledge (CSK), and domestic robotics. Key findings indicate a growing trend in integrating CSK into AI systems to enhance their decision-making capabilities, particularly in domestic environments. Studies such as "Commonsense Knowledge Extraction for Tidy-up Robotic Service in Domestic Environments" and "Large Language Models as Commonsense Knowledge" exemplify advancements in applying CSK for specific robotic tasks. Furthermore, the literature emphasizes the importance of Explainable AI

(XAI) in making AI systems more transparent and user-friendly, as highlighted in papers discussing the integration of CSK in human-robot collaborative tasks.

A significant theme identified in the literature is the challenge of effectively integrating CSK into robotic systems for practical applications. While there is a consensus on the potential of CSK to enhance robotic functionality, debates persist regarding the best methodologies and technologies for this integration. The literature also reveals a pattern of focusing on specific tasks or domains, with less emphasis on developing generalizable frameworks that apply CSK in dynamic domestic settings. Additionally, there is an ongoing discussion about balancing algorithmic efficiency with the need for user-centric design, particularly in the context of XAI.

The reviewed literature provides a solid foundation for this thesis by highlighting the current state and limitations of CSK integration in AI and domestic robotics. It sets the stage for the proposed research by underscoring the need for a more holistic, adaptable, and user-friendly approach to applying CSK in domestic robots. The Robo-CSK-Organizer, the focus of this thesis, aims to address these identified gaps by introducing a novel system that not only applies CSK for specific tasks but also adapts to a wide range of domestic environments. Moreover, by incorporating principles of XAI, the Robo-CSK-Organizer seeks to make AI decisions more transparent and understandable to users, thereby enhancing user trust and interaction quality.

2.9 Gap Analysis

Despite significant advancements in the integration of commonsense knowledge (CSK) in AI for domestic robotics, the literature reveals certain gaps. While existing studies like "Commonsense Knowledge Extraction for Tidy-up Robotic Service in Domestic Environments" and "Large Language Models as Commonsense Knowledge" focus on CSK application in specific contexts, there remains a need for comprehensive frameworks that can apply CSK in more generalized, dynamic domestic settings. Additionally, while some studies like "Robot Action Planning by Commonsense Knowledge in Human-Robot Collaborative Tasks" explore human-robot interactions, there is limited research on how CSK can enhance the adaptability of robots in varying and unforeseen domestic scenarios.

This research addresses these gaps by introducing the Robo-CSK-Organizer, a novel system that extends the application of CSK in AI for domestic robots. Unlike previous models, this system not only integrates CSK for specific tasks but also adapts to diverse and dynamic domestic environments, bridging the gap between static CSK application and real-world variability. Furthermore, this study contributes to the under-explored area of explainable AI (XAI) in domestic robotics, offering insights into how robots can communicate their decision-making processes in a user-friendly manner.

The Robo-CSK-Organizer is designed to dynamically apply CSK, considering spatial, temporal, and functional aspects of commonsense reasoning. This approach significantly enhances the robot's decision-making process, making it more adaptable and relevant to real-world domestic settings. Moreover, by incorporating XAI principles, the Robo-CSK-Organizer provides clearer and more understandable decision-making processes, thereby increasing user trust and interaction quality. Through a series of controlled experiments and comparative analyses with existing systems like the ChatGPT Organizer, this research demonstrates how the integration of advanced CSK and XAI can lead to more efficient, adaptable, and user-friendly domestic robotics, thus making a significant contribution to the field.

Chapter 3: Methodology

3.1 Proposed Approach Overview

This study introduces a novel methodological approach aimed at enhancing object organization in robotics through the integration of advanced AI technologies. By juxtaposing the newly developed Robo-CSK-Organizer with the ChatGPT Organizer, the research explores the frontier of automated decision-making in domestic robotics, charting a course towards more intelligent and adaptable robotic systems.

The comparative study between Robo-CSK and ChatGPT organizers is pivotal for advancing knowledge in AI-enhanced domestic robotics. This approach allows for a nuanced understanding of how different AI methodologies affect the efficiency and accuracy of object organization and classification in varied domestic environments. By comparing a rule-based, algorithmic approach (Robo-CSK) with a more flexible, AI-driven approach (ChatGPT), this study contributes to the ongoing discourse on optimizing AI for practical, everyday use in robotics. The findings from this comparison will offer valuable insights into how AI can adapt to real-world domestic settings, a key aspect of current AI research and development.

The methodology primarily aims to assess and compare the effectiveness of the Robo-CSK and ChatGPT organizers in context-sensitive object classification and placement. This comparative analysis is designed to reveal insights into the potential of AI-enhanced robotics in domestic settings, particularly in terms of decision-making accuracy, efficiency, and adaptability.

The Robo-CSK-Organizer and ChatGPT Organizer represent two distinct approaches to AI-driven object organization. The Robo-CSK-Organizer leverages a breadth-first search algorithm for decision-making, prioritizing context-relevant placement of objects. The ChatGPT Organizer, in contrast, employs a more flexible, query-based logic for object organization. The integration of BLIP and DETIC technologies is crucial, as they provide foundational support for room recognition and object identification.

The methodology is structured to facilitate a direct comparison between the two systems. This comparison is grounded in a set of well-defined metrics, including ambiguity resolution, consistency in object placement, and task relevance.

The methodological framework of this study is grounded in the principles of comparative analysis and experimental research in AI. Drawing on established models from the fields of machine learning and robotics, the methodology integrates and tests different AI approaches in a controlled, comparative setup. This framework aligns with the overarching research questions about the efficiency, accuracy, and adaptability of AI systems in domestic environments. By utilizing a structured comparative method, the study aims to generate empirical evidence that can inform future AI applications in domestic robotics, bridging a gap between theoretical AI models and their practical implementation.

The research design encompasses a thorough experimental setup, integrating BLIP and DETIC technologies with the robotic arm and ROS for precise object handling and classification. The comparison criteria are meticulously chosen to evaluate the efficacy of each system in real-world domestic scenarios.

3.2 Integration of NLP Models for Ambiguity Resolution

Natural Language Processing (NLP) models like FastText, Word2Vec, and GLOVE have been extensively used for semantic similarity analysis in various languages and contexts. FastText has been noted for its high accuracy in classification processes and is particularly effective due to its handling of out-of-vocabulary words through subword information, making it versatile across different datasets. Word2Vec and GLOVE are also prominent in embedding words into vectors, capturing semantic meanings and relationships. These models have been instrumental in tasks like sentiment analysis on social media platforms and other text analysis applications, where they are often combined with machine learning algorithms for enhanced performance [32][33].

The process involves converting words (objects and contexts) into vectors using these models. Each model embeds words in a high-dimensional space where semantically similar words are located closer to each other. The similarity between vectors (representing objects and contexts) is then computed, often using cosine similarity measures. This computation assesses how closely related an object is to a given context based on the model's learned semantic relationships. For instance, a study conducted to measure semantic similarity in Turkish texts utilized these pre-trained word embedding vectors, demonstrating their utility in understanding word-level semantic relationships [32][33]

The similarity scores generated by these models provide a quantifiable measure of semantic closeness between objects and contexts. By averaging the scores from FastText, Word2Vec, and GLOVE, a more robust and reliable semantic similarity score can be obtained. This averaged score serves as a ground truth for context assignment, enabling the determination of the most contextually appropriate location for an object. The higher the semantic similarity score between an object and a context, the more likely that context is the correct placement for the object. This methodology is crucial in scenarios where object-context associations are ambiguous, and a data-driven approach is needed to resolve these ambiguities [32][33].

3.3 Robo-CSK-Organizer Framework

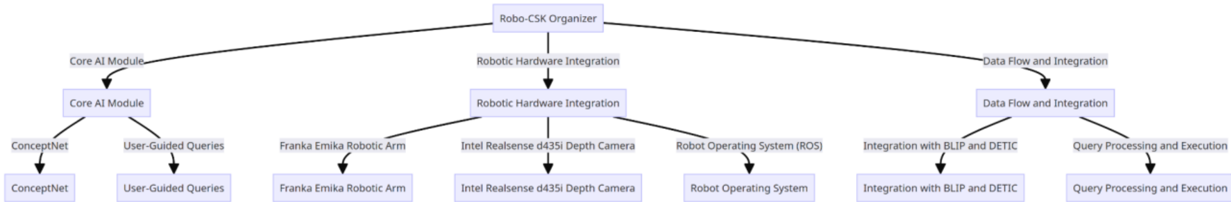


Figure 8: *Robo-CSK-Organizer Framework*

3.3.1 System Architecture **(See Figure 8)**

- **Command Terminal Interface:** The system is operated through command line inputs, allowing operators to initiate processes and monitor system performance.
- **Core AI Module:**
  - **Commonsense Knowledge (CSK) Processing:** Utilizes ConceptNet 5.5 to interpret and categorize objects based on contextual relevance.
  - **Breadth-First Search Algorithm:** The AI module employs this algorithm to identify optimal paths for object placement, integrating CSK to evaluate contextual relationships.

3.3.2 Robotic Hardware Integration

- **Franka Emika Robotic Arm:** The system features this state-of-the-art robotic arm, known for its precision and versatility in handling objects.
- **Intel Realsense d435i Depth Camera:** This camera provides crucial spatial and object data, crucial for accurate object detection and placement.
- **Robot Operating System (ROS):** Serves as the communication backbone, linking the AI module with the robotic arm and the camera. This setup ensures that the AI module's decisions are effectively translated into precise movements of the robotic arm.

3.3.3 Data Flow and Integration

- **Integration with BLIP and DETIC:** The system seamlessly integrates these technologies for efficient room recognition and object detection. BLIP identifies room contexts from images in the bins, while DETIC provides accurate object detection and localization.
- **Data Processing and Command Execution:** The AI module processes input data from both BLIP and DETIC, coupled with CSK, to make informed decisions about object placement. These decisions are then executed by the robotic arm under ROS control.

This architecture not only facilitates the technical functionality of the Robo-CSK-Organizer but also underscores its innovative approach to combining AI with robotics for enhanced object organization.

3.3.4 Functionality

The Robo-CSK-Organizer demonstrates advanced functionality in object organization by leveraging ConceptNet 5.5 for Commonsense Knowledge (CSK) processing. This is exemplified in the experimental setup where the system accurately categorized and placed objects like pears, remote controls, and computer mouse in contextually relevant bins representing different rooms like the kitchen, office, and bedroom.For instance, in one of the experiments, the system used the following logic paths, informed by CSK, for object placement:

The computer mouse was linked to the office through a path: Office (AtLocation | Weight: 4.90) <- computer (RelatedTo | Weight: 1.00) -> desktop (AtLocation | Weight: 1.00) <- computer_mouse | Degree of Separation: 3.

Similarly, the pear was associated with the kitchen: Kitchen (AtLocation | Weight: 7.21) <- food (RelatedTo | Weight: 2.32) -> apple (RelatedTo | Weight: 3.07) -> pear | Degree of Separation: 3.**(See figures 9a and 9b)**

These examples highlight the Robo-CSK-Organizer's ability to apply CSK for determining the most appropriate and logical location for objects, based on their conceptual and contextual relationships. This functionality underlines the system's innovative approach to enhancing decision-making in robotics through the integration of CSK.

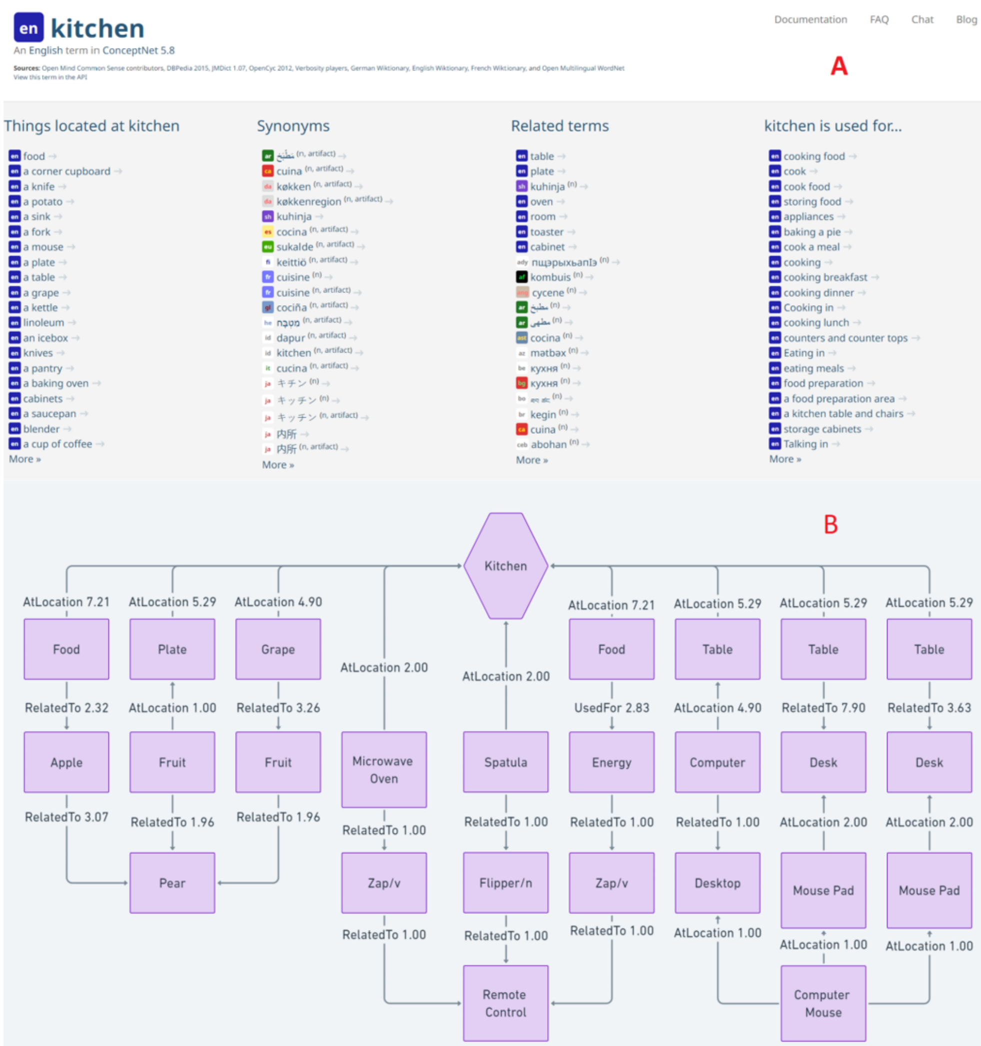


Figure 9a: *Searching the term Kitchen in ConceptNet Website*

Figure 9b: *Path diagram from Kitchen to Pear, Remote Control, and Computer*

3.4 ChatGPT Organizer Framework

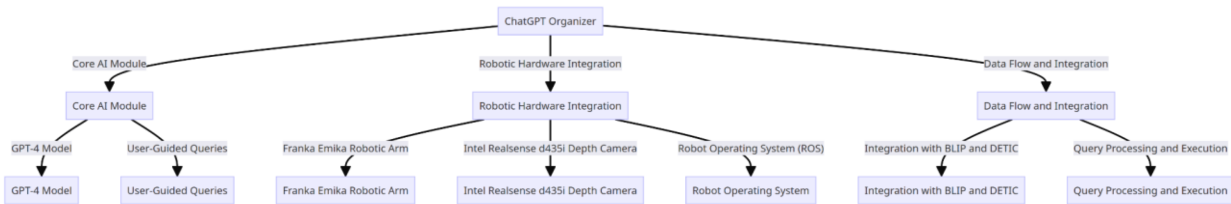


Figure 10: *Chat GPT Organizer Framework*

3.4.1 System Architecture **(See Figure 10)**

- **Command Terminal Interface:** Similar to the Robo-CSK-Organizer, the ChatGPT Organizer is controlled through command line inputs, facilitating real-time operational commands and system feedback.
- **Core AI Module:**
  - **Natural Language Processing (NLP):** Utilizes advanced NLP capabilities for interpreting and responding to user queries about object placement.
  - **Adaptive Decision-Making:** The AI module processes user queries and employs a flexible, context-aware logic to determine object placement.

3.4.2 Robotic Hardware Integration

- **Franka Emika Robotic Arm:** Employs the same robotic arm for object manipulation, ensuring precision and reliability.
- **Intel Realsense d435i Depth Camera:** Continues to provide vital spatial and object data for accurate object recognition and handling.
- **Robot Operating System (ROS):** Maintains the role of facilitating seamless communication between the AI module, the robotic arm, and the camera.

3.4.3 Data Flow and Integration

- **Integration with BLIP and DETIC:** These technologies are similarly integrated for room recognition and object detection, aiding the ChatGPT Organizer in making informed decisions.
- **Query Processing and Execution:** The AI module processes natural language queries from the operator, using the contextual information to guide the robotic arm in object placement, executed under ROS control.

This framework highlights the ChatGPT Organizer's unique approach in combining natural language processing with robotic control for object organization, offering a more interactive and user-guided experience.

3.4.4 Functionality

The ChatGPT Organizer utilizes a sophisticated approach combining OpenAI's GPT-4 model and user-guided queries for object organization. Unlike the Robo-CSK-Organizer's rule-based system, the ChatGPT Organizer dynamically processes natural language inputs to determine object placement.

For instance, if the system is presented with an object like "potato," it utilizes a Python script to communicate with the GPT-4 model. The operator inputs context preferences, and the model, considering these inputs, predicts the most appropriate context. In this case, the system might determine "kitchen" or "garden" as suitable placements for the potato, based on the given context options and tasks.

This interactive approach allows for a highly adaptable and user-responsive decision-making process, setting the ChatGPT Organizer apart in its ability to tailor object placement strategies to diverse and changing domestic environments.

3.5 Comparative Metrics: Ambiguity, Consistency, Task-Relevance and Explainability.

3.5.1 Metric Definition

- **Ambiguity:** Assesses the system's capacity to resolve scenarios where an object can logically belong to multiple contexts.
- **Consistency:** Measures the system's reliability in placing the same object in the same context under identical conditions.
- **Task Relevance:** Evaluates how effectively each system adapts object placement strategy to changing task priorities.
- **Explainability:** Gauges the ease with which users can understand and follow the logic behind each system's decisions, emphasizing the transparency and intelligibility of AI decision-making.

3.5.2 Measurement Techniques

- **Ambiguity:** Presenting both systems with objects that have multiple plausible placements and evaluating their decision-making.
- **Consistency:** Conducting repeated trials with identical objects and contexts to test decision-making consistency.
- **Task Relevance:** Observing system adjustments to classification decisions when task priorities change.
- **Explainability:** Analyzing the clarity and transparency of the reasoning process behind each system's object placement decisions.

Incorporating explainability as a metric is particularly significant, offering insights into the user-friendliness and trustworthiness of each system's decision-making process.

3.6 Implementation in Domestic Robotics

In the context of domestic robotics, the integration of BLIP and DETIC with the Franka Emika robotic arm and ROS represents a significant advancement in AI-enhanced household automation. This section outlines the technical workflow and real-world application of these systems in domestic environments.

Technical Workflow and System Coordination:

- BLIP and DETIC are crucial for room recognition and object identification. DETIC identifies objects, providing crucial data for the Robo-CSK and ChatGPT organizers to make decisions about object placement.
- The Franka Emika robotic arm, equipped with the Intel Realsense 3D infrared camera, acts upon these decisions, manipulating objects based on the AI's recommendations.
- ROS serves as the integrative platform, ensuring seamless communication and coordination between software (BLIP, DETIC) and hardware (robotic arm, camera). This integration allows for efficient execution of object organization tasks in real-time.

Application Context in Domestic Environments:

- These systems are designed for use in various domestic settings, like kitchens, living rooms, and bedrooms, where they can assist in organizing everyday items, contributing to household management and efficiency.

Challenges and Solutions:

- One of the main challenges is ensuring accurate and consistent object recognition and categorization in diverse and dynamic home environments. This is addressed through the continuous refinement of AI algorithms and sensory inputs.

- Another challenge is the seamless integration of software and hardware components to work in unison, which is tackled through rigorous testing and optimization within the ROS framework.

Real-World Relevance:

- The implementation of these technologies in domestic robotics showcases the potential of AI to enhance daily living, offering practical solutions for object organization and task management.
- These systems also serve as a step towards solving the "opaque-box" problem in AI, as discussed in "Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence" [4], by providing more transparent and understandable AI decision-making processes in household contexts.

This implementation holds promise for making domestic robots more intuitive, responsive, and useful for a wide range of household tasks, significantly contributing to the field of home automation and robotics.

3.7 Data Collection and Analysis Methods

3.7.1 Data Sources

The primary data sources for this research comprise an extensive object database and pre-defined contextual scenarios, encompassing a wide array of everyday objects and potential environments. This database is integral for providing a diverse and comprehensive set for experimentation.

3.7.2 Collection Process

Data collection occurs in real-time during experimental execution. This process involves recording decisions made by the Robo-CSK and ChatGPT organizers for object placement in various contexts. The system logs, encompassing object identification, placement decisions, and contextual information, are methodically captured for subsequent analysis. Additionally, NLP models such as FastText, Word2Vec, and GLOVE are implemented to calculate semantic similarity scores, crucial in establishing a baseline for ambiguity and consistency assessments.

3.7.3 Analysis Techniques

Quantitative Analysis: Utilizes Python's SciPy and Pandas libraries for statistical analysis, focusing on accuracy, consistency, and adaptability metrics. The semantic similarity scores derived from the NLP models play a critical role here. They provide a quantifiable measure of the relationship between objects and contexts, aiding in evaluating ambiguity and consistency across different scenarios.

Qualitative Analysis: Involves scrutinizing decision logs and system outputs to assess explainability and decision-making logic. This analysis is particularly important for understanding the reasoning behind each system's decisions, especially in cases where these decisions might appear counterintuitive or complex.

Comparative Analysis: Both systems are compared to evaluate their relative effectiveness in varying scenarios. This includes analyzing the performance of the Robo-CSK Organizer, which integrates Detic, BLIP, and ConceptNet for object detection and contextual understanding, against the ChatGPT Organizer's deep learning-based decision-making process. Integration of NLP Models:

The integration of FastText, Word2Vec, and GLOVE models significantly contributes to data analysis, particularly in measuring ambiguity and consistency. These models are used to compute similarity scores between objects and their potential contexts. By averaging these scores, a more comprehensive understanding of each object's contextual relevance is achieved, serving as a benchmark for comparing the organizers' decisions.

This comprehensive data collection and analysis approach ensures a thorough understanding of each system's capabilities and limitations. It significantly contributes to advancing AI-driven domestic robotics, providing insights into the efficacy of integrating different AI models and techniques in real-world applications.

3.8 Validation of Methods

Reliability and Accuracy:

- The reliability and accuracy of the methods will be ensured through rigorous testing and calibration of the systems before actual data collection. This includes verifying the precision of the object detection and context recognition algorithms (BLIP and DETIC) and ensuring the robotic arm's accurate response to AI decisions. Regular system checks will be conducted to maintain consistency in performance.

Testing Protocols:

- A series of pre-defined testing protocols will be employed, involving controlled experiments with a set range of objects and contexts. These protocols include redundancy checks, system response time measurements, and accuracy assessments against established benchmarks.

Evaluation Criteria:

- The success and validity of the methodology will be evaluated based on specific criteria, including the systems' accuracy in object placement, consistency across multiple trials, adaptability in varying contexts, and the clarity of decision-making processes. Statistical significance and practical applicability in domestic environments will also form crucial aspects of the evaluation process.

This methodological validation ensures that the research results are reliable, accurate, and applicable to real-world scenarios in domestic robotics.

3.9 Ethical Considerations

Ethical Concerns:

- Potential ethical concerns in this research primarily revolve around the implications of human-robot interactions, particularly regarding privacy, autonomy, and reliance on AI in domestic settings. The research also considers the potential biases in AI decision-making, especially in the context of object categorization and placement.

Mitigation Strategies:

- To address these concerns, the research incorporates rigorous data privacy protocols, ensuring no sensitive personal data is collected or stored. Measures are taken to ensure the AI systems do not infringe on user autonomy or promote excessive dependence on robotic assistance. Bias mitigation strategies include diverse data inputs and continuous monitoring for biased patterns in decision-making.

Compliance with Standards:

- The research methodology strictly adheres to the established ethical standards in robotics and AI research. This includes compliance with guidelines on responsible AI use, transparency in AI decision-making processes, and consideration of long-term societal impacts of deploying such technologies in domestic environments.

These considerations ensure that the research is ethically sound and aligned with the broader objectives of responsible AI development and application.

Chapter 4: Experimental Evaluation

4.1 Experimental Setup

The experiments were designed to assess the performance of Robo-CSK-Organizer and ChatGPT Organizer in contexts requiring ambiguity resolution, consistency, and task-relevance adaptability. The test objects were categorized into groups like Fruits, Vegetables, Beverages, and Snacks, and were evaluated across different contexts such as kitchen, office, playroom, etc. The experimental setup involved multiple trials for each scenario to ensure robustness in the results.

4.2 Experiment 1: Resolving Ambiguity

The primary objective of this experiment was to assess the capability of Robo-CSK-Organizer and ChatGPT Organizer in effectively resolving ambiguous contexts associated with various objects. The test objects were selected from a diverse range of categories such as Fruits, Vegetables, and Office Supplies. Each object was evaluated against a set of predefined contexts (e.g., kitchen, office, playroom), and the organizers' proficiency in identifying the most contextually relevant setting for each object was critically analyzed.

A crucial aspect was the implementation of FastText along with Word2Vec and GloVe to enhance the robustness of semantic similarity analysis. The FastText model was specifically chosen for its ability to generate word embeddings for out-of-vocabulary (OOV) words by considering subword information, making it a valuable asset in handling a wide range of objects and contexts.

The procedure involved calculating similarity scores for each object-context pair using FastText, Word2Vec, and GloVe models. The average of these scores provided a comprehensive view of the semantic relationship between the object and the context. This methodological approach ensured a more nuanced and accurate assessment of ambiguity, considering the strengths of each model. (**See Figure 11**)

To quantify ambiguity, the experiment employed a novel approach that measured the smallest difference in average similarity scores across various contexts for each object. This method involved analyzing the semantic similarity scores derived from the FastText, Word2Vec, and GloVe models. The closer the average of these scores were for two different contexts, the higher the ambiguity associated with that object-context pair.

The ambiguity resolution capability was then tested by choosing two contexts for each object that were considered ambiguous. Essentially the models determined that the object could be sorted in either of the two concepts.

	object	category	context	fasttext_similarity	word2vec_similarity	glove_similarity	average_similarity
177	adhesive_tape	Office Supplies	garden	0.067916	0.037400	-0.026665	0.026217
172	adhesive_tape	Office Supplies	playroom	0.070947	0.031446	-0.006632	0.031920
171	adhesive_tape	Office Supplies	office	0.075511	0.045630	0.098610	0.073250
176	adhesive_tape	Office Supplies	pantry	0.072087	0.125592	0.050121	0.082600
174	adhesive_tape	Office Supplies	bedroom	0.024524	0.111487	0.143355	0.093122
170	adhesive_tape	Office Supplies	kitchen	0.062532	0.131583	0.118485	0.104200
179	adhesive_tape	Office Supplies	bathroom	0.054122	0.249015	0.198313	0.167150
173	adhesive_tape	Office Supplies	living_room	0.415202	0.082728	0.124765	0.207565
175	adhesive_tape	Office Supplies	dining_room	0.504062	0.063728	0.057654	0.208481
178	adhesive_tape	Office Supplies	laundry_room	0.561421	0.162899	0.175983	0.300101

Figure 11: *Semantic scores between object-context pairs.*

Then a "ground_truth was established by taking the higher of the two similarity scores. The responses of the Robo-CSK-Organizer and the Chat GPT organizer were then compared against the 'ground truth' context. This comparison provided a clear understanding of each system's ability to navigate and resolve ambiguity in object-context associations.

The results from this experiment are seen in the **(See figure 12)**. The accuracy of each system was calculated as the percentage of instances where the system's chosen context matched the ground truth context. In total, 45 objects were analyzed. The Robo-CSK-Organizer correctly organized 19 of these objects against the ground truth observations established by the three NLP models, while the ChatGPT organizer correctly organized 24 of these objects against ground truth observations.

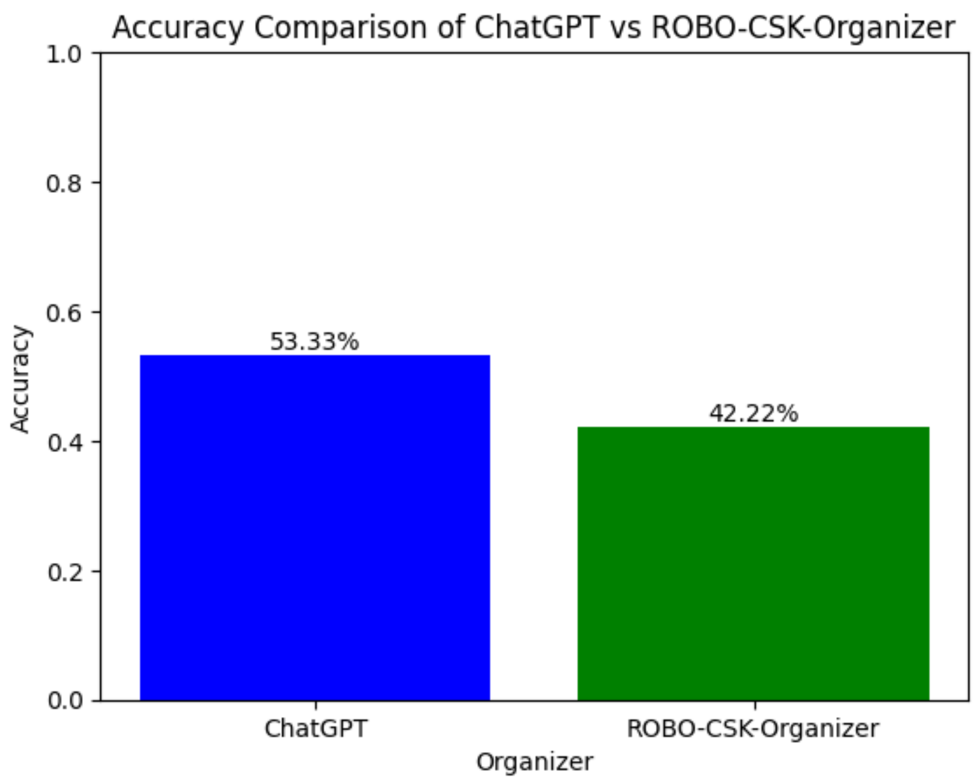


Figure 12: *Ambiguity resolution between ChatGPT and Robo-CSK-Organizer*

4.3 Experiment 2: Ensuring Consistency

This experiment was designed to evaluate the consistency of the Robo-CSK-Organizer and ChatGPT Organizer. The aim was to determine if these systems provided stable and repeatable results when faced with identical queries over multiple iterations. Consistency in this context refers to the systems' ability to repeatedly choose the same context for an object across numerous trials, an essential trait for reliable knowledge organization systems.

4.3.1 Methodology

Object Selection: A variety of objects were selected from diverse categories including Personal Items, Clothing, Office Supplies, and Toys. This range ensured a broad evaluation spectrum.

Context Groups: Specific context groups, namely culinary locations (kitchen, garden, pantry, dining room), were chosen for the evaluation. These contexts were relevant to the selected object categories and provided a controlled environment for testing.

Test Iterations: Each object was queried against the systems ten times, with the systems being asked to organize the object into one of the provided contexts.

Data Collection: The responses from both systems were recorded for each iteration. The most frequently chosen context by each system was then identified as the predominant choice.

4.3.2 Results Analysis

Data Visualization: The consistency of each system was visually represented using bar graphs (**See Figure 13)** , comparing the frequency of the most common response for each object across both systems.

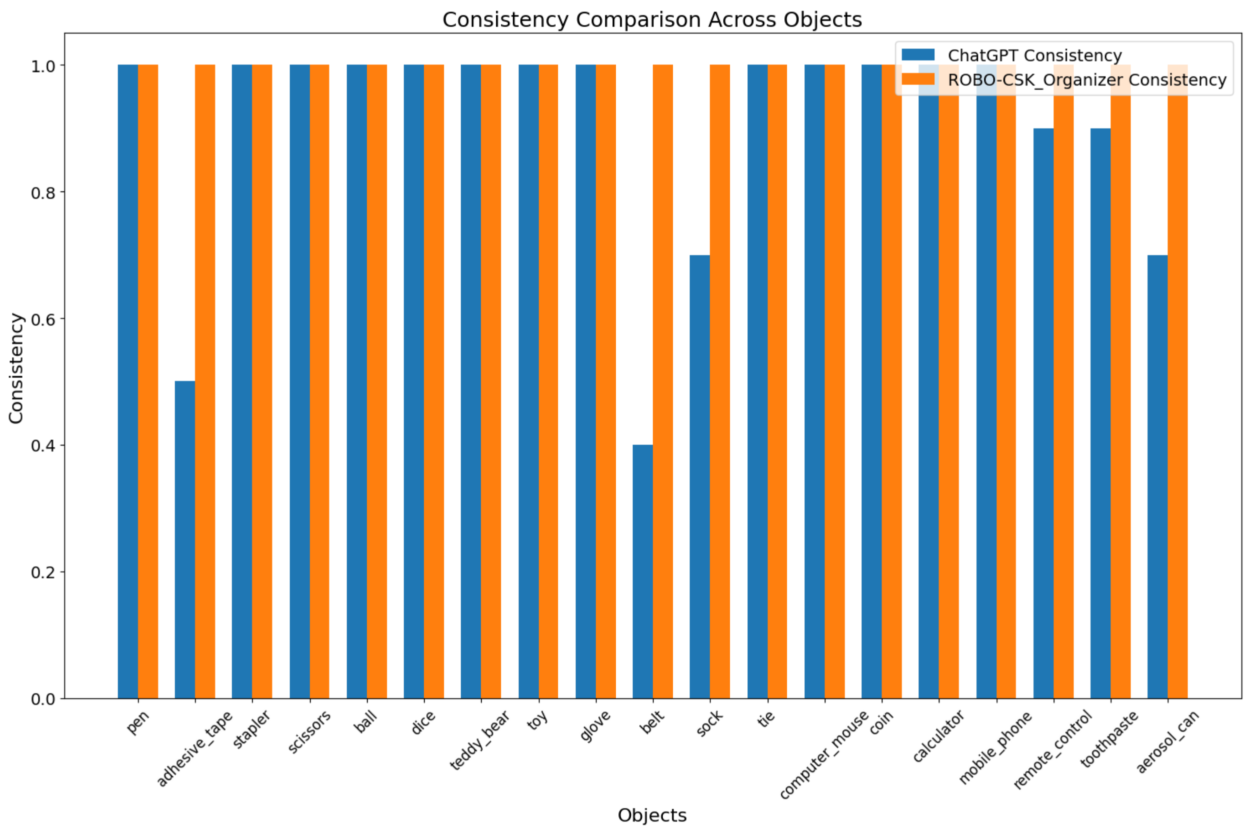


Figure 13: *Robo-CSK-Organizer vs Chat GPT in Consistency*

Consistency Measurement: The consistency score was computed as the ratio of the number of times the most common response was given to the total number of trials. This provided a quantifiable measure of how often each system stuck to its most preferred choice.

4.3.3 Findings

Variability in Responses: The experiment highlighted the degree of consistency (or lack thereof) in each system's contextual organization. For some objects, a high level of consistency was observed, while for others, responses varied more significantly. As predicted, the Robo-CSK-Organizer had a 100 % consistency rate across all object/location pairs. This is mostly due to the nature of conceptnet being a static knowledge graph. As can be seen from the **figure 13,** the Chat GPT implementation was not as consistent as the Robo-CSK-Organizer for all objects, specifically for adhesive tape, belt, sock, remote_control, toothpaste, and aerosol can.

4.4 Experiment 3: Task-Relevance Adaptability

The objective of this experiment was to assess the adaptability of the Robo-CSK-Organizer and ChatGPT Organizer in response to varying context-specific directives. This adaptability is crucial for applications where context relevance can frequently shift. The experiment sought to understand if the systems could recalibrate their responses when directed to focus on alternative contexts, differing from their initial preferences.

4.4.1 Experimental Design

Baseline Assessment: The experiment commenced with a baseline assessment where 'apple' was queried against four contexts (kitchen, living room, bedroom, bathroom) without any focus directives. This was done to determine the systems' natural inclination or preference for context association.

Adaptability Testing: Post the baseline test, the systems were then prompted to focus on the remaining three contexts, one at a time. This was to observe if the systems could adapt their responses when a specific context was emphasized.

Data Collection: For both baseline and adaptability tests, each context query was repeated ten times to ensure robust data collection. The responses were compiled into respective dataframes for detailed analysis.

Focus Contexts: The focus contexts for the adaptability test were selected based on the baseline preference. The context which emerged as the most preferred in the baseline was excluded in the adaptability test, emphasizing the remaining contexts.

4.4.2 Analysis Methodology

Response Counting: The number of times each context was chosen by the systems in both baseline and adaptability phases was tallied.

Visualization: Bar graphs were employed to visually represent the frequency of responses for each context in both test phases. This provided a clear comparative view of the systems' response patterns.

Comparative Analysis: The adaptability of each system was gauged by comparing the most common responses from the baseline test with those from the adaptability test.

4.4.3 Findings

Baseline Preferences: The experiment identified kitchen to be a clear baseline preference for sorting Apple. This was true for both the Chat GPT implementation and the Robo-CSK-implementations.

Response Shifts: In the adaptability phase, the chat gpt implementation did not cause any observable shifts, despite telling Chat GPT to focus on the contexts in question. However, there were observable shifts on the Robo-CSK-Implementation on one of the locations, which was the bedroom. When it came to sorting Apple, having the Robo-CSK-Detector focus on the bedroom prompted it to sort the apple in the bedroom. See **Figure 14 and 15** for details. The reasoning paths the Robo-CSK-Organizer opted for when sorting to kitchen, and then sorting to bedroom once given the bedroom focus, is given below.

Path: kitchen (AtLocation | Weight: 7.21) <- food (RelatedTo | Weight: 2.32) -> apple

Path: bedroom (AtLocation) <- house (AtLocation) <- apple

Also included are real world examples of the Robo-CSK Organizer organizing a bottle into a dining room prior to becoming focused, and then organizing into the laundry room after being prompted to focus as a laundry room robot. Please See **figures 16 and 17** for the images of the study, along with their respective logical paths.

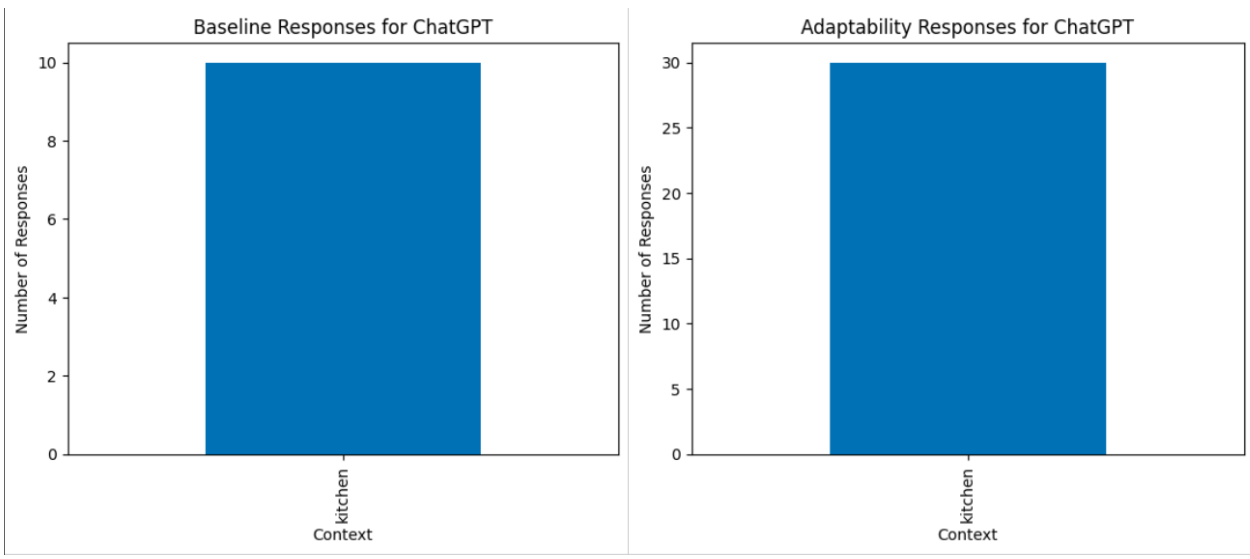


Figure 14: *Baseline vs Adaptability responses for ChatGPT*

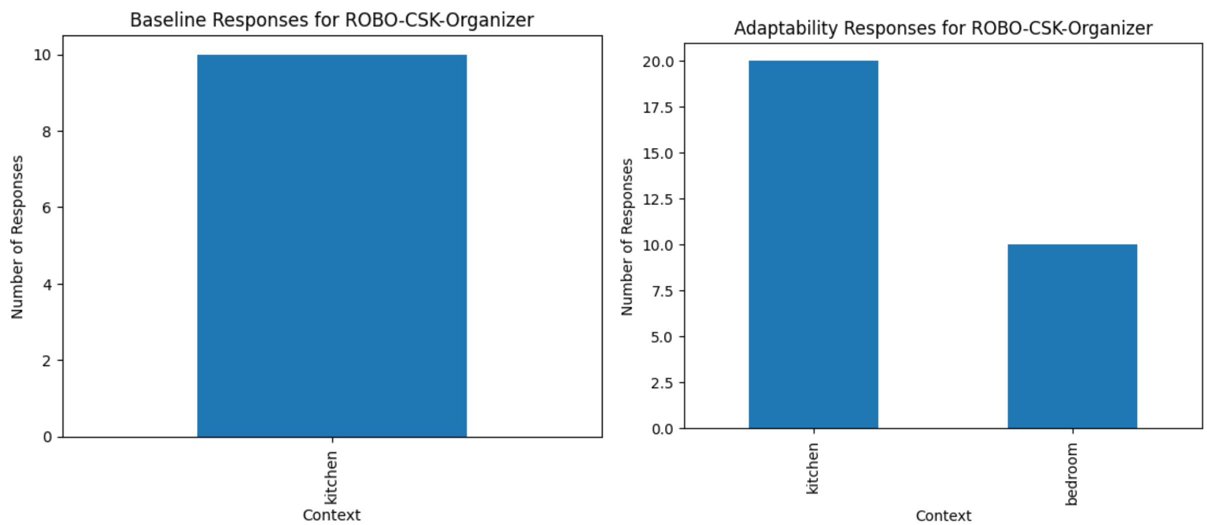


Figure 15: *Baseline vs Adaptability responses for Robo-CSK-Organizer*

Figure 16: *Robo-CSK-Organizer Unfocused Bottle Placement*
Note: Bottle into dining room ***Contexts in images are:*** *(Playroom, Dining Room, Laundry Room, Living Room)* Logic Path is Dining Room (AtLocation) < glass (RelatedTo) > container (IsA) < bottle

Figure 17: *Robo-CSK-Organizer Laundry Room Focused Bottle Placement*
Note: bottle into laundry room ***Contexts in images are:*** *(Playroom, Dining Room, Laundry Room, Living Room)* Logic Path is Laundry Room (AtLocation) < sink (AtLocation) > hair (RelatedTo) < bottle

4.5 Experiment 4: Explainability

The experiment on explainability aimed to evaluate the capacity of Robo-CSK-Organizer and ChatGPT Organizer to not only make decisions but also elucidate the reasoning behind these decisions. This aspect is crucial, especially in scenarios where the decision might appear counterintuitive or potentially hazardous.

Robo-CSK-Organizer's Explainability: The Robo-CSK-Organizer utilizes a combination of Detic for object detection, BLIP for context recognition, and most importantly, ConceptNet for commonsense knowledge (CSK) to infer object locations. This system, by leveraging CSK, can provide a logical path or reasoning behind its decisions. For instance, in the case where beer was incorrectly sorted into the playroom, the Robo-CSK-Organizer provided a clear logical path: playroom (UsedFor) -> fun (RelatedTo) -> party (RelatedTo) -> beer. **(See**

Figure 19). This example demonstrates how the system connects the concepts of playrooms, fun, parties, and beer, leading to its conclusion.

ChatGPT Organizer's Opaque Decision-Making: In contrast, ChatGPT, as a deep learning-based model, operates more like an "opaque box", making it challenging to understand the reasoning behind its decisions. An example of this is ChatGPT's decision to sort scissors into the playroom, which is potentially dangerous for children. Without an explainable AI framework, it is difficult to ascertain the logic or reasoning process that led to this decision, making it harder to adjust or correct the model's behavior.**(See Figure 18)**

Importance of Explainability: The key difference here is that while both systems may make similar errors, the Robo-CSK-Organizer's explainability feature allows for a deeper understanding and subsequent correction of these errors. Explainability, thus, is not just about making the right decisions but also about understanding and learning from the wrong ones. This becomes particularly important in robotic systems where precision and safety are paramount. See **figures 19 and 20** for how Chat GPT and Robo-CSK-Organizer have opted to sort scissors.

Object	**Chat GPT Context**	**Robo-CSK-Organizer Context**	**Robo-CSK-Organizer Path**
beer	living_room	playroom	Playroom (UsedFor) -> fun (RelatedTo) -> party (RelatedTo) -> beer
scissors	playroom	dining_room	dining_room(AtLocation) <- table (RelatedTo) -> desk (AtLocation) <- scissors

Figure 18: *Logic of putting Scissors and Beer into erroneous contexts.*

Figure 19: *Robo-CSK-Organizer Placement of scissors into Dining Room*

Figure 20: *ChatGPT-Organizer Placement of scissors into Child's Playroom*

4.6 Comparative Analysis of Robo-CSK-Organizer vs. ChatGPT Organizer

In the comparative analysis of the Robo-CSK-Organizer and the ChatGPT Organizer, each system's performance was evaluated against four critical metrics: ambiguity resolution, consistency in object placement, task-relevance adaptability, and explainability.

4.6.1 Ambiguity Resolution:

- Robo-CSK-Organizer: Utilizes ConceptNet, Detic, and BLIP technologies, enabling it to understand environmental contexts effectively. This helps in correctly identifying the most logical context for an object. However, it occasionally struggled with ambiguities inherent in real-world scenarios.
- ChatGPT Organizer: Employs ChatGPT for commonsense reasoning combined with the same Robo-CSK-Organizer object recognition technologies. It showed a higher proficiency in resolving ambiguities compared to the Robo-CSK-Organizer. Its integration of language processing allowed it to interpret user queries and context more dynamically.

4.6.2 Consistency in Object Placement:

- Robo-CSK-Organizer: Achieved a 100% consistency rate in object placement across repeated tasks. Its reliance on ConceptNet's structured knowledge base contributed to this high level of consistency.
- ChatGPT Organizer: Demonstrated less consistency, especially with varied objects. This variability is possibly due to its flexible, context-aware logic which, while adaptable, might lead to different conclusions under similar conditions.

4.6.3 Task-Relevance Adaptability:

- Robo-CSK-Organizer: Exhibited significant adaptability in adjusting classification logic based on the changing nature of tasks. This adaptability is crucial in dynamic domestic environments where task priorities can shift rapidly.
- ChatGPT Organizer: While adaptable, it was less effective than the Robo-CSK-Organizer in recalibrating its decision-making processes in response to changing tasks. This might be attributed to the inherent complexities of natural language processing and the challenges of integrating it with object recognition.

4.6.4 Explainability:

- Robo-CSK-Organizer: Offered superior explainability, providing clear and logical reasoning behind its decisions. This clarity is likely due to its structured approach to knowledge organization and decision-making.
- ChatGPT Organizer: Faced challenges in achieving the same level of explainability as the Robo-CSK-Organizer. The 'opaque box' nature of deep learning models in language processing can make it difficult to trace the exact reasoning behind decisions.

While both systems have their strengths, the Robo-CSK-Organizer generally outperforms the ChatGPT Organizer in terms of consistency, adaptability and explainability. However, the ChatGPT Organizer shows a stronger performance in ambiguity resolution. This comparative analysis underscores the importance of a balanced approach in AI system design, blending structured knowledge bases with advanced language processing and object recognition technologies.

4.7 Discussion of Results

The experimental results offer significant insights into the evolving landscape of AI in domestic settings, specifically through the analysis of the Robo-CSK-Organizer and the ChatGPT Organizer. These findings have broader implications for the development of AI systems, particularly in their application within home environments.

4.7.1 Ambiguity Resolution:

- ChatGPT Organizer's Edge: The superior performance of the ChatGPT Organizer in ambiguity resolution suggests an advancing trend in AI towards more nuanced semantic understanding. Its success in these scenarios indicates the potential for future AI systems to better interpret and respond to the complexities of human environments.
- Robo-CSK-Organizer's Potential: The effectiveness of the Robo-CSK-Organizer, though slightly less in accuracy, highlights the importance of integrating structured knowledge bases in AI for domestic use. This approach could be pivotal in developing systems that comprehend and interact with everyday life nuances more effectively.

4.7.2 Consistency:

- Benchmarking Reliability: The perfect score in consistency by the Robo-CSK-Organizer sets a new benchmark for reliability in AI applications. This consistency is crucial for fostering trust and predictability in AI systems, making them more viable and acceptable for everyday household use.

- Real-World Application Significance: The findings emphasize the need for AI systems in domestic settings to provide stable and repeatable outcomes, underscoring the value of predictability in fostering user trust and acceptance.

4.7.3 Adaptability:

- Dynamic Decision-Making: The pronounced adaptability of the Robo-CSK-Organizer in task-relevant classifications reflects a significant stride towards creating AI systems that can dynamically adjust to changing household needs and preferences.
- Future of Flexible AI: This adaptability underlines the potential for future AI systems to be more contextually aware and flexible, adapting to the varied and unpredictable nature of domestic environments.

4.7.4 Explainability:

- Transparent AI: The Robo-CSK-Organizer's advantage in explainability points towards an emerging need in AI development: creating systems that are not only efficient but also transparent in their decision-making processes. This transparency is key in domestic settings where users must understand and trust AI decisions.
- Challenges in Deep Learning Models: The difficulties faced by the ChatGPT Organizer in explainability shed light on the ongoing challenges in making deep learning models more interpretable and user-friendly.

4.7.5 Integration of Advanced Technologies:

- Elevating AI Capabilities: The integration of Detic and BLIP technologies with CSK frameworks in the Robo-CSK-Organizer represents a significant advancement in the field. This synthesis of technologies enhances adaptability, consistency, and explainability, setting a new standard in AI system design.
- Future Directions: These results underscore the necessity for continued innovation in AI, particularly in combining advanced technologies with user-centric design principles. The Robo-CSK-Organizer exemplifies a balanced approach, paving the way for intelligent, adaptable, and user-friendly domestic robotics.

Overall, these findings highlight the importance of developing AI systems that are not only task-efficient but also understand the subtleties of human environments, maintain consistency, and offer clear explainability. The comparative study of the Robo-CSK-Organizer and ChatGPT Organizer provides valuable lessons for future developments in AI, suggesting a shift towards more integrated, user-centered AI systems in domestic robotics.

Chapter 5: Conclusions and Future Work

5.1 Summary of Findings

The thesis provides comprehensive insights into the integration of Commonsense Knowledge (CSK) in AI, with a specific focus on domestic robotics. Key findings from the study include:

- **Integration of CSK in AI:** The Robo-CSK-Organizer system successfully demonstrated the integration of CSK into AI for domestic robotics. This integration was achieved through the utilization of advanced technologies such as ConceptNet, Detic, and BLIP. These technologies enabled the Robo-CSK-Organizer to effectively classify and place objects in contextually appropriate settings.
- **Comparative Analysis:** A detailed comparative analysis was conducted between the Robo-CSK-Organizer and the ChatGPT Organizer. This analysis revealed the strengths and weaknesses of each system in various performance metrics, including ambiguity resolution, consistency, task-relevance adaptability, and explainability.
- **Performance Metrics:**
    - **Ambiguity Resolution:** The ChatGPT Organizer displayed higher proficiency in ambiguity resolution compared to the Robo-CSK-Organizer. This was attributed to its dynamic language processing capabilities, which allowed for a more nuanced interpretation of user queries and contexts.
    - **Consistency in Object Placement:** The Robo-CSK-Organizer achieved a 100% consistency rate in object placement, benefiting from ConceptNet's structured

knowledge base. In contrast, the ChatGPT Organizer showed less consistency, likely due to its flexible, context-aware logic.

- **Task-Relevance Adaptability:** The Robo-CSK-Organizer exhibited significant adaptability in adjusting classification logic based on changing task priorities. This adaptability was crucial in dynamic domestic environments where task priorities can shift rapidly. The ChatGPT Organizer, while adaptable, was less effective in recalibrating its decision-making processes in response to changing tasks.
- **Explainability:** The Robo-CSK-Organizer offered superior explainability, providing clear and logical reasoning behind its decisions. This clarity was likely due to its structured approach to knowledge organization and decision-making. The ChatGPT Organizer faced challenges in achieving the same level of explainability, as the 'opaque box' nature of deep learning models in language processing can make it difficult to trace the exact reasoning behind decisions.

The findings underscore the importance of integrating CSK into AI systems for domestic robotics, highlighting the potential for such systems to be more intuitive, consistent, and adaptable, with a clear emphasis on explainability.

5.2 Limitations of the Study

5.2.1 Scope of Experiments:

- **Limited Object and Context Selection:** The experiments were confined to a pre-defined set of objects and contexts, which may not encompass the full range of scenarios encountered in real-world domestic settings. This limitation could affect the applicability of the findings to more diverse or unpredictable home environments.

- **Controlled Environment:** The experimental setup was conducted in controlled conditions that may not accurately represent the dynamic and sometimes chaotic nature of real households. Consequently, the performance of the systems in actual domestic scenarios might vary.

5.2.2 Model Generalizability:

- **Specific AI Models and Technologies:** The study primarily focused on the Robo-CSK-Organizer and the ChatGPT Organizer, which were built on particular AI models and technologies. This specificity may limit the generalizability of the findings to other types of domestic robots or AI systems that use different technologies or have different operational frameworks.
- **Adaptability to Diverse Environments:** The study's findings might not directly apply to all types of domestic environments, especially those with unique characteristics or non-standard layouts. The generalizability of the results to such diverse settings remains uncertain.

5.2.3 Additional Limitations:

- **Human Interaction Dynamics:** The study did not fully explore the dynamics of human-robot interaction, particularly how users might adapt to or influence the behavior of the robots. The variability in human behavior and preferences could significantly impact the effectiveness of the systems in practical use.

- **Long-Term Reliability:** The experiments did not address the long-term reliability and maintenance requirements of the systems. Over time, wear and tear, software updates, or changes in the domestic environment could affect the systems' performance.
- **Data and Privacy Concerns:** Although not explicitly mentioned in the thesis, a potential limitation relates to data privacy and security. The use of AI in domestic settings raises concerns about the collection, storage, and use of personal data, which must be considered in the broader application of such technologies.

In recognizing these limitations, the study underscores the need for cautious interpretation of the findings and suggests avenues for further research to address these gaps.

5.3 Recommendations for Future Work

5.3.1 Broader Data Sets:

- **Diversified Object Range:** Expand the range of objects used in experiments to include a more diverse array of household items, including those with less clear-cut classifications or those specific to certain cultures or lifestyles.
- **Varied Contextual Scenarios:** Incorporate a wider variety of scenarios that better represent the complexity and unpredictability of real-world environments, including unstructured and dynamically changing domestic settings.

5.3.2 Enhanced Explainability:

- **Advanced Explainability Techniques:** Develop methods to enhance the transparency of AI decision-making processes, particularly for deep learning models. This could involve integrating techniques from the field of explainable AI (XAI) to make the logic behind decisions more accessible and understandable to users.

- **User-Centric Explanations:** Focus on creating explanations that are tailored to the end-user's perspective, ensuring that they are easily comprehensible without requiring technical expertise.

5.3.3 Cross-Context Application:

- **Beyond Domestic Environments:** Explore the application of the Robo-CSK-Organizer and similar systems in non-domestic settings, such as offices, healthcare facilities, or public spaces, to assess their versatility and adaptability in different contexts.
- **Cross-Cultural Adaptability:** Investigate the systems' performance in diverse cultural settings, considering variations in household structure, lifestyle, and object usage.

5.3.4 User-Centered Design:

- **Interactive and Intuitive Interfaces:** Improve the user interfaces of these systems to make them more interactive and intuitive, focusing on ease of use for non-technical users.
- **User Feedback Integration:** Implement mechanisms for collecting and incorporating user feedback into the system's decision-making process, enhancing its relevance and user satisfaction.
- **Personalization Features:** Develop personalization features that allow the system to adapt to individual user preferences and habits over time, making it more aligned with the specific needs and routines of its users.

These recommendations aim to broaden the scope and enhance the effectiveness of AI systems in domestic and other environments, focusing on user-centric design and adaptability to diverse contexts and user needs.

5.4 Implications of the Research

5.4.1 Advancement in AI-Driven Robotics:

- **Integration of CSK:** The study demonstrates the effective integration of Commonsense Knowledge (CSK) into AI, significantly advancing the field of domestic robotics. This integration enhances the ability of robots to understand and interact with the complexities of human environments.
- **Robotic Decision-Making:** The findings highlight the importance of adaptability and explainability in AI-driven robotic systems, contributing to the development of more sophisticated decision-making algorithms.
- **Benchmark for Future Systems:** The research sets a benchmark for future AI systems in domestic robotics, showing the potential of combining advanced technologies like ConceptNet, Detic, and BLIP for practical applications.

5.4.2 Practical Applications:

- **Enhanced User Experience:** The study's findings have the potential to influence the development of domestic robots that are more intuitive and adaptable, significantly improving user experience in household settings.
- **Real-World Application**: The research underscores the practicality of AI in everyday life, paving the way for robots that can more effectively assist with domestic tasks, enhance personal convenience, and provide support in various household activities.

5.4.3 Theoretical Contributions:

- **Framework for CSK Integration:** The research offers valuable insights into how CSK can be integrated into AI systems, providing a theoretical framework that can guide future advancements in AI and robotics.

- **Influence on AI Development:** The study contributes to the theoretical understanding of AI, particularly in terms of processing and interpreting complex information akin to human cognition, which is crucial for the evolution of AI technologies.
- **Cross-Disciplinary Insights:** The findings have implications beyond robotics, offering insights into fields like natural language processing, machine learning, and human-computer interaction, potentially influencing a wide range of AI applications.

The implications of this research are far-reaching, with the potential to influence both practical applications and theoretical advancements in AI and robotics, particularly in the realm of domestic settings. The study's contributions extend beyond the technical aspects, offering a comprehensive understanding of how AI can be designed to be more aligned with human needs and environments.

5.5 Contribution to the Field

5.5.1 Innovative Integration of Technologies:

5.5.2 Unique Combination of CSK and AI:

- **Pioneering Effort:** The Robo-CSK-Organizer is a trailblazer in melding Commonsense Knowledge (CSK) with cutting-edge AI technologies. This unique combination is a significant stride forward, revolutionizing the way robots interpret and engage with their surroundings.
- **Unique Approach in Robotics:** This integration redefines robotic perception and interaction in domestic environments, enabling robots to understand and respond to human spaces more intuitively.

5.5.3 Utilization of Advanced Technologies:

- **Synergistic Technology Integration:** The system's adept use of ConceptNet, Detic, and BLIP exemplifies the power of integrating diverse technologies. This synergy yields more nuanced object classification and a deeper understanding of context.
- High Degree of Autonomy:
  - **Detic's Role:** Specializing in object recognition, Detic empowers the system to accurately identify a wide array of household items.
  - **BLIP's Contextual Insight:** BLIP's ability to discern different environments enhances the system's contextual awareness, crucial for appropriate interaction.
  - **Guidance by ConceptNet:** Leveraging ConceptNet's extensive commonsense knowledge, the system can intelligently determine the most suitable placements or uses for objects in given contexts.
- **Efficient and Self-Reliant System:** This harmonious integration results in a highly efficient, autonomous system, adept at navigating domestic settings in a manner that resonates with human logic and preferences.
- **Precedent for Future AI Systems:** Such an integration sets a new standard for future AI systems, emphasizing technological convergence and enhanced capabilities.

5.5.4 Enhancing AI System Transparency:

- **Advancements in Explainable AI (XAI):** The research contributes significantly to the field of XAI, showcasing how AI systems, particularly those used in domestic settings, can be designed to be more transparent and understandable in their decision-making processes.

- **User-Friendly AI Decisions:** By focusing on explainability, the study addresses a critical aspect of AI technology, ensuring that users can understand and trust the decisions made by AI systems. This is particularly important in domestic environments where non-technical users interact with AI systems.
- **Setting Standards for Future Developments:** The study establishes new standards for transparency and user interaction in AI systems. This contribution is likely to influence the direction of future research and development in AI, particularly in terms of making complex technologies more accessible and comprehensible to the general public.

The contributions of this research are multifaceted, encompassing both technological innovation and advancements in AI system transparency. By showcasing the practical application of these concepts in domestic robotics, the study not only advances the field technologically but also makes significant strides in addressing the challenges of user interaction and system transparency in AI.

5.6 Reflections on the Research Process

5.6.1 Learning from Challenges:

- **Handling Ambiguity in Object Contexts:** A significant challenge was the ambiguity inherent in categorizing objects for different contexts. For instance, determining whether a 'spoon' should be placed in the kitchen or a dining area required nuanced understanding. Addressing this ambiguity using Fasttext, GLOVE, and Word2Vec improved our approach to resolving ambiguity, particularly in enhancing the accuracy of context-specific classifications.

- Integration of Advanced Technologies: One challenge was the effective integration of advanced technologies like Detic, BLIP, and ConceptNet. Each technology serves a specific purpose, and the challenge lay in harmonizing these different functionalities to create a cohesive system. This required careful design and testing to ensure that the individual strengths of each technology were effectively utilized and that they worked seamlessly together.
- Handling Real-World Variability: Another significant challenge was designing the system to handle the variability and unpredictability of real-world domestic environments. This included recognizing and appropriately categorizing a wide range of objects in constantly changing settings. Addressing this challenge involved developing algorithms capable of adapting to new and unforeseen situations, enhancing the system's flexibility and robustness.
- Balancing User Interaction and Autonomy: Striking the right balance between user interaction and the autonomy of the AI system was a complex challenge. The system needed to be sufficiently autonomous to make smart decisions independently while remaining intuitive and user-friendly, allowing users to interact with and guide it when necessary.
- Ethical Considerations in AI Development: Ethical considerations, such as privacy concerns and the potential for AI bias, posed another challenge. The research process involved not only developing effective AI solutions but also ensuring that these solutions were ethically sound and aligned with societal values and norms.
- Technical Limitations and Scalability: Overcoming technical limitations, such as computational constraints and ensuring the scalability of the system for broader

application, was a challenge. This involved optimizing algorithms for efficiency and ensuring that the system could be scaled up to accommodate a larger variety of tasks and environments without a loss in performance.

- Data Collection and Management: Collecting and managing a diverse and extensive dataset for training and testing the system was also challenging. The system required a vast array of data to learn from, necessitating efficient data collection, storage, and processing mechanisms.

These challenges contributed significantly to the learning process and the overall development of the research, providing valuable insights into the complexities of AI applications in domestic robotics. Addressing these challenges also led to a deeper understanding of the various aspects that need to be considered when developing sophisticated AI systems for real-world applications.

5.6.2 Methodological Considerations:

- **Comparative Analysis Between Systems:** The research involved a detailed comparative analysis of the Robo-CSK-Organizer and the ChatGPT Organizer. This method was crucial in understanding how each system processed information and made decisions. For example, the comparative analysis helped identify that the Robo-CSK-Organizer excelled in consistency and explainability due to its structured knowledge base, while the ChatGPT Organizer was better at resolving ambiguous contexts due to its advanced language processing capabilities.
- **Effectiveness of Experimental Design**: The experimental design included controlled tests where both systems were presented with an array of objects in varying contexts.

This approach was effective in quantifying each system's ability to correctly categorize objects and adapt to changing scenarios. However, the experiments also highlighted the need for more dynamic testing environments that better mimic real-world conditions.

Reflecting on these aspects of the research process reveals the depth of learning and adaptation required in AI and robotics research. The challenges faced and the methodologies employed not only enriched the research outcomes but also provided a clearer direction for future studies in this evolving field.

Works Cited

[1]Z. Zhao, W. Lee, and D. Hsu, “Large Language Models as Commonsense Knowledge for Large-Scale Task Planning,” Oct. 2023.

[2]G. Cui, W. Shuai, and X. Chen, “Semantic Task Planning for Service Robots in Open Worlds,” *Future Internet*, vol. 13, no. 2, p. 49, Feb. 2021, doi: https://doi.org/10.3390/fi13020049.

[3]M. Sridharan and T. Mota, “Towards combining commonsense reasoning and knowledge acquisition to guide deep learning,” *Autonomous Agents and Multi-Agent Systems*, vol. 37, no. 1, Nov. 2022, doi: https://doi.org/10.1007/s10458-022-09584-4.

[4]C. Zednik, “Solving the Black Box Problem: a Normative Framework for Explainable Artificial Intelligence,” *Philosophy & Technology*, vol. 34, no. 2, pp. 265–288, Dec. 2019, doi: https://doi.org/10.1007/s13347-019-00382-7.

[5]V. K. Chaudhri *et al.*, “Knowledge graphs: Introduction, history, and Perspectives,” *AI Magazine*, vol. 43, no. 1, pp. 17–29, Mar. 2022, doi: https://doi.org/10.1002/aaai.12033.

[6]X. Jiang *et al.*, “On the Evolution of Knowledge Graphs: a Survey and Perspective,” Arxiv, Oct. 2023.

[7]H. Liu and P. Singh, “ConceptNet — a Practical Commonsense Reasoning Tool-Kit,” *BT Technology Journal*, vol. 22, no. 4, pp. 211–226, Oct. 2004, doi: https://doi.org/10.1023/b:bttj.0000047600.45421.6d.

[8]R. Speer, J. Chin, and C. Havasi, “ConceptNet 5.5: an Open Multilingual Graph of General Knowledge,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, Feb. 2017, doi: https://doi.org/10.1609/aaai.v31i1.11164.

[9]N. Tandon, A. S. Varde, and G. de Melo, “Commonsense Knowledge in Machine Intelligence,” *ACM SIGMOD Record*, vol. 46, no. 4, pp. 49–52, Feb. 2018, doi: https://doi.org/10.1145/3186549.3186562.

[10]M. Sap, V. Shwartz, A. Bosselut, Y. Choi, and D. Roth, “Introductory Tutorial: Commonsense Reasoning for Natural Language Processing,” *Association for Computational Linguistics*, pp. 27–33, 2020, doi: https://doi.org/10.18653/v1/P17.

[11]T. Young, E. Cambria, I. Chaturvedi, H. Zhou, S. Biswas, and M. Huang, “Augmenting End-to-End Dialogue Systems with Commonsense Knowledge,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, Apr. 2018, doi: https://doi.org/10.1609/aaai.v32i1.11923.

[12]C. Richardson and L. Heck, “Commonsense Reasoning for Conversational AI: A Survey of the State of the Art,” Arxiv, Feb. 2023.

[13]A. Holzinger, P. Kieseberg, E. Weippl, and A. M. Tjoa, “Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction: From Machine Learning to Explainable AI,” *Lecture Notes in Computer Science*, pp. 1–8, 2018, doi: https://doi.org/10.1007/978-3-319-99740-7_1.

[14]S. Storks, S. Edu, Q. Gao, and J. Chai, “RECENT ADVANCES IN NATURAL LANGUAGE INFERENCE: A SURVEY Recent Advances in Natural Language Inference: A Survey of Benchmarks, Resources, and Approaches,” Arxiv, Feb. 2020.

[15]N. Bian, X. Han, Le Sun, H. Lin, Y. Lu, and B. He, “ChatGPT is a Knowledgeable but Inexperienced Solver: An Investigation of Commonsense Problem in Large Language Models,” Arxiv, Mar. 2023.

[16]P. P. Ray, “ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope,” *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 121–154, 2023, doi: https://doi.org/10.1016/j.iotcps.2023.04.003.

[17]T. B. Sheridan, “Human–Robot Interaction,” *Human Factors: the Journal of the Human Factors and Ergonomics Society*, vol. 58, no. 4, pp. 525–532, Apr. 2016, doi: https://doi.org/10.1177/0018720816644364.

[18]M. A. Goodrich and A. C. Schultz, “Human-Robot Interaction: a Survey,” *Foundations and Trends® in Human-Computer Interaction*, vol. 1, no. 3, pp. 203–275, 2007, doi: https://doi.org/10.1561/1100000005.

[19]M. Callegari *et al.*, “Tools and Methods for Human Robot Collaboration: Case Studies at i-LABS,” *Machines*, vol. 10, no. 11, p. 997, Oct. 2022, doi: https://doi.org/10.3390/machines10110997.

[20]A. Ostrowski *et al.*, “Ethics, Equity, & Justice in Human-Robot Interaction: A Review and Future Directions,” Sep. 2022. Accessed: Nov. 29, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9900805

[21]X. Gao, R. Gong, Y. Zhao, S. Wang, T. Shu, and S.-C. Zhu, “Joint Mind Modeling for Explanation Generation in Complex Human-Robot Collaborative Tasks,” Arxiv, Jul. 2020. Accessed: Nov. 29, 2023. [Online]. Available: https://arxiv.org/abs/2007.12803

[22]E. Prassler and K. Kosuge, “Domestic Robotics,” *Springer Handbook of Robotics*, pp. 1253–1281, 2008, doi: https://doi.org/10.1007/978-3-540-30301-5_55.

[23]J. Zhong, C. Ling, A. Cangelosi, A. Lotfi, and X. Liu, “On the Gap between Domestic Robotic Applications and Computational Intelligence,” *Electronics*, vol. 10, no. 7, p. 793, Mar. 2021, doi: https://doi.org/10.3390/electronics10070793.

[24]Y. Li, “Research and Application of Deep Learning in Image Recognition,” *2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA)*, Jan. 2022, doi: https://doi.org/10.1109/icpeca53709.2022.9718847.

[25]M. Feng, “Research on the Application of Deep Learning in Computer Image Recognition,” *Journal of Physics: Conference Series*, vol. 1915, no. 4, p. 042034, May 2021, doi: https://doi.org/10.1088/1742-6596/1915/4/042034.

[26]J. Kumar, P. Bedi, S. Goyal, and X. Zhao, “Research and application of deep learning in image recognition” *Journal of Physics: Conference Series PAPER • OPEN ACCESS Journal of Physics: Conference Series*, vol. 2425, p. 12047, 2023, doi: https://doi.org/10.1088/1742-6596/2425/1/012047.

[27]X. Zhou, R. Girdhar, A. Joulin, P. Krähenbühl, and I. Misra, “Detecting Twenty-thousand Classes using Image-level Supervision,” Arxiv, Jul. 2022. Accessed: Nov. 30, 2023. [Online]. Available: https://arxiv.org/abs/2201.02605

[28]J. Li, D. Li, C. Xiong, and S. Hoi, “BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation,” Arxiv, Feb. 2022. Accessed: Nov. 30, 2023. [Online]. Available: https://arxiv.org/abs/2201.12086

[29]Weerachai Skulkittiyut *et al.*, “Commonsense knowledge extraction for Tidy-up robotic service in domestic environments,” *IEEE*, Nov. 2013, doi: https://doi.org/10.1109/arso.2013.6705507.

[30]L. Kunze, M. Tenorth, and M. Beetz, “Putting People’s Common Sense into Knowledge Bases of Household Robots,” Sep. 2010.

[31]C. Conti, A. Varde, and W. Wang, “Robot Action Planning by Commonsense Knowledge in Human-Robot Collaborative Tasks,” IEEE, Oct. 2020. Accessed: Nov. 30, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9216410

[32]E. Dharma, F. Lumban Gaol, H. Leslie, H. Warnars, and B. Soewito, “THE ACCURACY COMPARISON AMONG WORD2VEC, GLOVE, AND FASTTEXT TOWARDS CONVOLUTION NEURAL NETWORK (CNN) TEXT CLASSIFICATION,” *Journal of Theoretical and Applied Information Technology*, vol. 31, no. 2, 2022, Available: https://www.jatit.org/volumes/Vol100No2/5Vol100No2.pdf

[33]C. Wang, P. Nulty, and D. Lillis, “A Comparative Study on Word Embeddings in Deep Learning for Text Classification,” *Proceedings of the 4th International Conference on Natural Language Processing and Information Retrieval*, Dec. 2020, doi: https://doi.org/10.1145/3443279.3443304.

[34]P. Persaud, A. S. Varde, and S. A. Robila, “Enhancing Autonomous Vehicles with Commonsense: Smart Mobility in Smart Cities,” *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)* , pp. 1008–1012, Nov. 2017, doi: https://doi.org/10.1109/ictai.2017.00155.

[35]A. Garg, N. Tandon, and A. S. Varde, “I Am Guessing You Can’t Recognize This: Generating Adversarial Images for Object Detection Using Spatial Commonsense (Student Abstract),” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 10, pp. 13789–13790, Apr. 2020, doi: https://doi.org/10.1609/aaai.v34i10.7166.

[36]A. Varde and G. Weikum, “Common Sense Knowledge in Domain-Specific Knowledge Bases,” Max-Planck-Institut fur Informatik, (Saarbrücken, Germany), Aug. 2015. Accessed: Dec. 07, 2023. [Online]. Available: https://web.cs.wpi.edu/~aparna/MPII_2015_TechReportCSKdomainKB.pdf

[37]C. J. Conti, A. S. Varde, and W. Wang, “Human-Robot Collaboration With Commonsense Reasoning in Smart Manufacturing Contexts,” *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 1784–1797, Jul. 2022, doi: https://doi.org/10.1109/tase.2022.3159595.

[38]L. Paulino, C. Hannum, A. S. Varde, and C. J. Conti, “Search Methods in Motion Planning for Mobile Robots,” *Intelligent Systems and Applications: Proceedings of the 2021 Intelligent Systems Conference (IntelliSys)*, vol. 3, pp. 802–822, Aug. 2021, doi: https://doi.org/10.1007/978-3-030-82199-9_54.

[39]C. J. Conti, A. S. Varde, and W. Wang, “Task quality optimization in collaborative robotics,” *2020 IEEE International Conference on Big Data (big data)*, pp. 5652–5654, Dec. 2020, doi: https://doi.org/10.1109/bigdata50022.2020.9378498.

[40]T.-P. Nguyen, S. Razniewski, A. Varde, and G. Weikum, “Extracting Cultural Commonsense Knowledge at Scale,” *arXiv.org*, Apr. 30, 2023. https://arxiv.org/abs/2210.07763

[41]C. J. Conti, A. S. Varde, and W. Wang, “Web Perspectives in Robotics Applications,” *ACM SIGWEB Newsletter*, vol. 2023, no. Winter, pp. 1–22, Mar. 2023, doi: https://doi.org/10.1145/3583849.3583853.

[42]I. Chernyavsky, A. S. Varde, and S. Razniewski, “CSK-Detector: Commonsense in object detection,” *2022 IEEE International Conference on Big Data (Big Data)*, Dec. 2022, doi: https://doi.org/10.1109/bigdata55660.2022.10020915.

[43]A. Garg, N. Tandon, and A. Varde, “CSK-SNIFFER: Commonsense Knowledge for Sniffing Object Detection Errors,” Apr. 2022. Accessed: Dec. 07, 2023. [Online]. Available: https://ceur-ws.org/Vol-3135/bigvis_short2.pdf

[44]S. Razniewski, N. Tandon, and A. S. Varde, “Information to Wisdom: Commonsense Knowledge Extraction and Compilation,” *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pp. 1143–1146, Mar. 2021, doi: https://doi.org/10.1145/3437963.3441664.

[45]R. Hidalgo, A. DeVito, N. Salah, A. S. Varde, and R. W. Meredith, “Inferring Phylogenetic Relationships using the Smith-Waterman Algorithm and Hierarchical Clustering,” *2022 IEEE International Conference on Big Data (Big Data)*, Dec. 2022, doi: https://doi.org/10.1109/bigdata55660.2022.10020454.

[46]R. Hidalgo, N. Salah, R. Jetty, and A. Jetty, “Personalizing Text-to-Image Diffusion Models by Fine-Tuning Classification for AI Applications,” Sep. 2023.

Appendices

A. Algorithms and Code-Snippets

Algorithm 1: Robo-CSK-Detector Main Algorithm

Input: Video feed, ConceptNet knowledge base.

Output: Objects in appropriate contexts.

```
Initialize robot vision, Detectron2.
Scan bins, recognize contexts using BLIP, store in CSV
for each video frame do
        Detect and label objects with Detectron2.
        for each object do
                Query ConceptNet for object's context.
                if context matches bin's context then
                    Place object in matched bin.
                else
                Move on to next object.
        end if
  end for
end for
Optional: Display annotated video frame.
```

Algorithm 2 BFS through ConceptNet.

Input: OBJECTS list, LOCATIONS list, API_END-POINT, cached data (if available)
Output: Paths between objects and locations.

```
Load cached data if available.
function fetch related data(node, rel type)
    Retrieve and cache data from API ENDPOINT.
end function
function bfs modified with pathwise visited(start)
    Initialize queue with start node and related data.
    while queue is not empty do
        Dequeue node and related data.
        Determine relationship types based on node's degree.
        for each relation type do
            Fetch related data.
            Process fetched data to determine next nodes and paths.
            Enqueue valid next nodes.
        end for
    end while
    return paths.
end function
Initialize all paths as empty dictionary.
for each location in LOCATIONS do
    Call bfs modified with pathwise visited and store paths.
end for
Save all paths to 'paths modified.json'.
Print summary results.
```

Code Snippet 1: Extracting Relevant Context for an Object.

Input: All_Paths.json, contexts.csv, object
Output: Most relevant context for the object and its path.

```
# Load the data
with open('All_Paths.json', 'r') as file:
    data = json.load(file)
# Read the CSV file
contexts_csv = '/home/parronj1/Rafael/local_detectron/contexts.csv'
df = pd.read_csv(contexts_csv)
# Extract detected contexts
desired_contexts = []
desired_contexts.extend(df["Context"])
# Get the most relevant context for the object
context, path, weight, degree_of_separation = get_object_context(data, object, desired_contexts)
print(f"The most relevant context for {object} is {context}.")
# Check if a relevant context is found
if context and not context.startswith("No paths found"):
    print(f"Path: {path}")
else:
    print("No relevant paths found.")
```

B. Links to project related videos

https://drive.google.com/drive/folders/1Sd6Yjt3rP-ytNrMYFAXVDhNYqGOzbsMy?usp=drive_link

https://drive.google.com/drive/folders/1JB4qc6r5uo8_o62wUmvCH0cYE1A9_Kpq?usp=drive_link

https://drive.google.com/drive/folders/1wTTaPqMLUJg51b3nXgDrT5RH3ISF-eDq?usp=drive_link