Universität Siegen / Cognizant Mobility

Utilizing large language models and prompt-engineering driven intent taxonomies for effective user intent detection in Conversational AI system.

Piyush Sham Patil
1634587
26th March 2025

Examiners:
Prof. Dr. Markus Cristinziani
Dr. Christian Weber

www.uni-siegen.de

Agenda

Introduction

Literature Review

Taxonomy of Intents

Implementation

Evaluation Methodology

Results

Conclusion

1

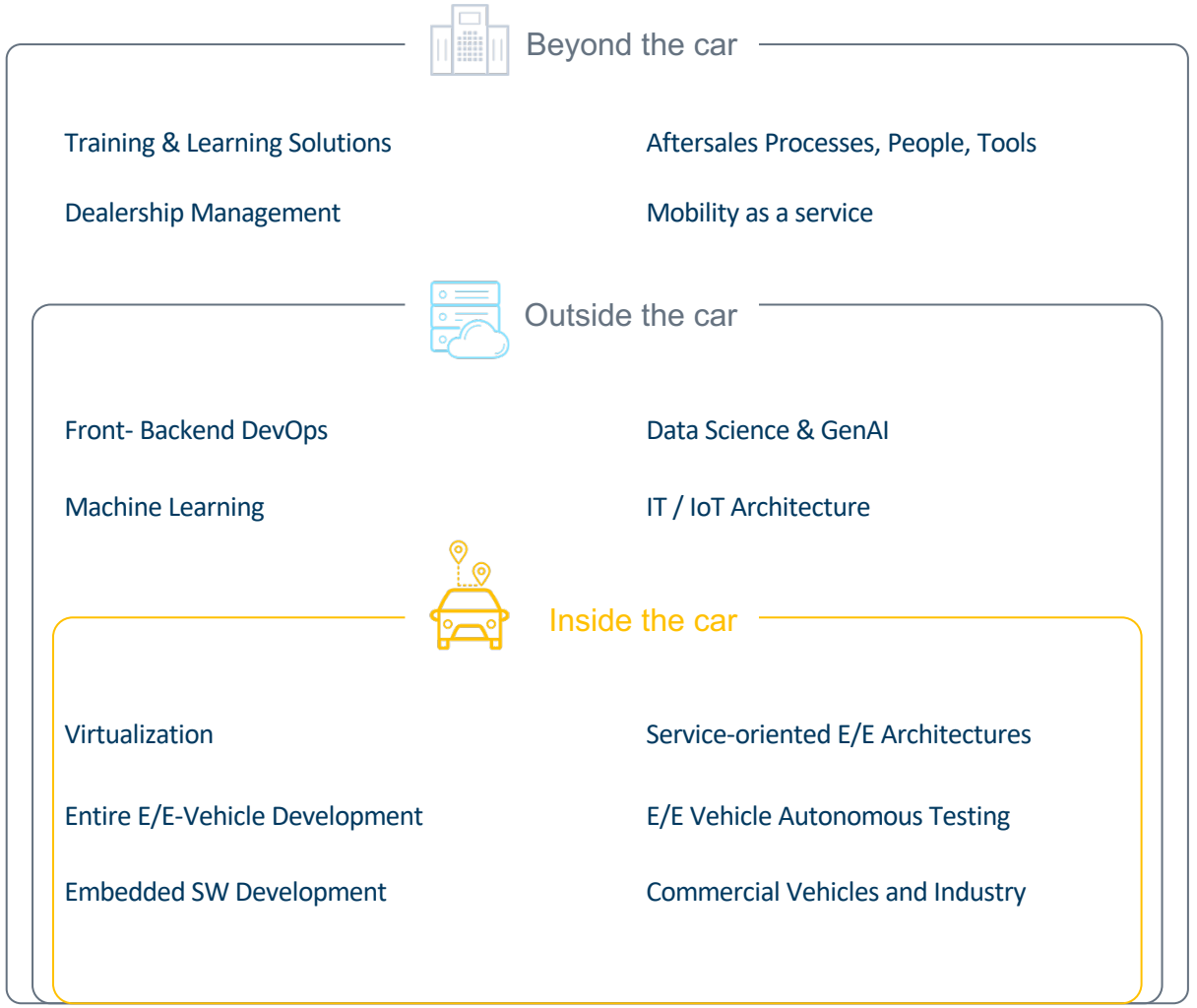
Introduction

Cognizant Mobility GmbH

- This work was conducted in partnership with Cognizant Mobility GmbH, Munich, Germany, a company specializing in advanced automotive vehicle solutions.

- **Cognizant Mobility**: A leading provider of advanced electronic and IT solutions for emerging market trends.

- Expertise includes:
  - Software development and operations.
  - Connectivity solutions.
  - End-to-end vehicle E/E development.
  - Data analytics and cybersecurity.
  - System integration and testing.



Cognizant Mobility

Motivation

- Intent is central to effective communication, representing the purpose or goal behind actions or queries.
- Accurate intent classification is vital for AI systems to ensure effective and personalized user responses.
- Enhance intent recognition in AI systems to improve user interactions and AI-driven experiences.
- **AI Systems:**
  - **VERA**: Very Enhanced Road Assistant, unveiled at IAA 2023, serves as a travel companion for passengers in autonomous vehicles.
  - **Onboarding AI Assistant**: Supports new employees by addressing HR queries and enhancing experience.

Problem Statement

- Current gaps in VERA and Onboarding AI Assistant:
  - Lack robust mechanisms for detecting user intent.
  - Misinterpreting intent leads to irrelevant or unsatisfactory responses.
  - Impacts responsiveness and user experience negatively.

- Need for advanced techniques like **Prompt Engineering:**
  - Improve intent classification.
  - Enhance query reformulation for better response retrieval.

Research Questions

1. Optimizing intent classification.
    - How can user inputs and prompts be optimized for accurate intent detection?
    - Focus: Refining interaction through Prompt Engineering.

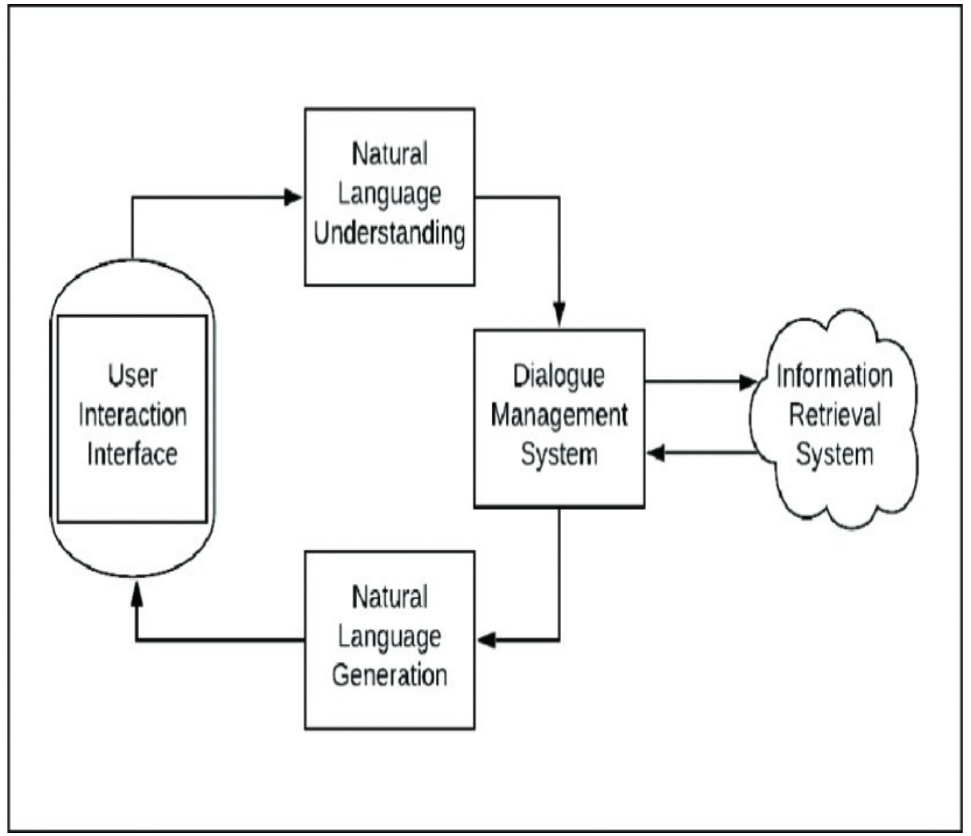
2. Validating intent classification accuracy.
    - How can intent classification be validated in conversational AI systems?
    - Enhance query reformulation for better response retrieval.

2

Literature Review

Natural Language Processing Fundamentals

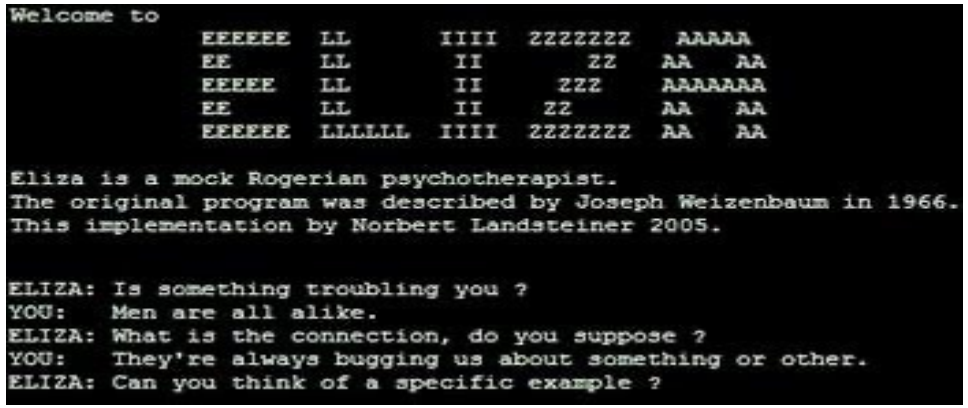
- NLP is a fundamental component of Conversational AI, consisting of two branches:
  - **Natural Language Understanding (NLU):**
    - Interprets the meaning behind user inputs.
  - **Natural Language Generation (NLG):**
    - Transforms structured data into human-like, coherent responses.
    - Bridges the gap between machine understanding and human communication.
  - **Dialogue Management System (DMS):**
    - Handles conversational flow and determines AI's next actions.
    - Manages state transitions to maintain context during conversations.



Advancements in Conversational AI

- Conversational AI is a branch of Artificial Intelligence that simulates and automates speech- and text-based interactions.

- Origins of Conversational AI: Development of ELIZA by Joseph Weizenbaum at MIT in 1960s. Simulated human conversation.



OpenAI

- **Modern Conversational AI (2024)**: Sophisticated models like OpenAI's GPT-4. Generates human-level responses.

- **Driving Factors for Advancement**: **Machine Learning (ML)** and **Deep Learning (DL)** improved accuracy and performance of models.

Intent Recognition Overview

- **Traditional Machine Learning:**
  - **Models:** Random Forests, Support Vector Machines (SVM).
  - **Strengths:** Achieved up to 75% precision in early experiments.
  - **Challenges:** Reliance on extensive labeled datasets, Limited scalability for real-world applications.

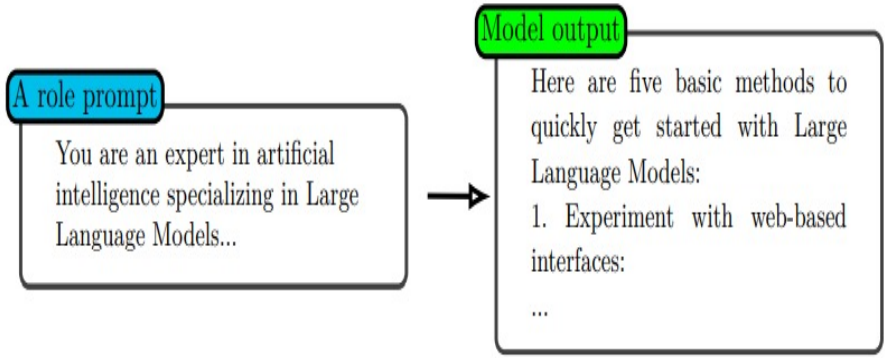
- **Deep Learning Advancements:**
  - **Models:** Recurrent Neural Networks (RNNs), Bidirectional Long Short-term Memory (bi-LSTMs).
  - **Strengths:** Sequential dependency capture and feature extraction.
  - **Challenges:** High computational demands, misclassification in long queries.

- **Large Language Models (LLMs):**
  - Transformative shift in intent recognition.
  - **Models:** Generative pre-trained transformer – 4 (GPT-4).
  - **Capabilities:** Human-like natural language understanding, context-aware query reformulation.
  - **Advantages:** Reduced reliance on labeled datasets, handling of nuanced and domain-specific queries.

Prompt Engineering Techniques

- **Prompt Engineering:** Intentional creation and fine-tuning of input prompts to guide LLMs toward precise, relevant, and cohesive responses.

- **Importance**:
  - Refines model outputs for user-specific needs.
  - Optimizes response quality and aligns with task objectives.



- **Core Strategies**:
  - **Clarity and Specificity**: Craft detailed and unambiguous prompts to enhance relevance.
  - **Role Prompting**: Assign specific roles (e.g., "expert," "assistant") to refine context and output.
  - **Prompt Chaining**: Sequential prompts build structured, multi-step reasoning.
  - **Chain-of-Thought Prompting (CoT)**: Integrates logical steps for complex reasoning tasks.

Taxonomy of Intents

Taxonomies for Intent Recognition

- **Purpose of Taxonomies**:
  - Enhance understanding of user queries in conversational AI.
  - Enable accurate intent classification for diverse use cases.

- **Granularity of Taxonomy**:
  - Includes 15 sub-intents across 5 primary categories.
  - Designed for precise interpretation and classification of user queries.

- **Five Primary Intent Categories**:
  - Informational Intent
  - Advisory Intent
  - Navigational Intent
  - Administrative Intent
  - Personal Interaction Intent

Informational Intent

- Descriptive Intent 📄
- Procedural Intent 🛠️
- Exploratory Intent 🔍

Advisory Intent

- Analytical Intent 📊
- Operational Intent ✅
- Consultative Intent 🗣️

Navigational Intent

- Location-Based Intent 📍
- Feature Navigation Intent 🧭
- Service Navigation Intent 🧮

Administrative Intent

- Scheduling Intent ⏰
- Compliance Intent ⚖️
- Access Management Intent 🔑

Personal Interaction Intent

- Conversational Engagement 💬
- Decision Support 🧠
- Reflection and Insights 🌟

Reformulation Template in Intent Recognition

- **Purpose of Reformulation:**
  - Improves query clarity and aligns it with the intended information retrieval task.
  - Generates two responses: original query vs. reformulated query (for testing and evaluation).

- **Challenges with User Queries:**
  - Lack of specificity or precision may hinder accurate responses.
  - Users struggle with terminology or framing the question effectively.

- **Utility and Adaptability:**
  - Templates are modular and adaptable for various research contexts.
  - Contributes to both practical and theoretical advancements in conversational AI.

Descriptive Intent

Provide/Offer/Explain/List a detailed overview/key features/significance of the topic or subject, specifically focusing on the specific aspect or feature.

Procedural Intent

Provide/Explain/Describe the step-by-step process for completing the task, including any necessary conditions/requirements and specific details relevant to the task.

Exploratory Intent

Provide/Explain/Describe a comprehensive overview of the topic/concept, focusing on key aspects/significance/relevant details.

Analytical Intent

Compare/Evaluate/Analyse two or more items, focusing on specific criteria or aspects. Which option offers the best/most comprehensive/preferred outcome?

Operational Intent

What/Which/Can you recommend actionable steps, best practices, recommended options for a specific task or decision to achieve a desired outcome or goal?

4

Implementation

Prompt Chaining

- Sequential processing of user queries through specialized tasks.
- Utilizes LangChain framework to create independent LLM modules for specific tasks.
- **Key Chains in the System:**
  - **Label Generation Chain**: Extracts object-topic pairs from raw user queries.
  - **Refinement Chain:** Enhances object-topic pairs with external contextual information.
  - **Intent Detection Chain:** Maps refined object-topic pair and query to a predefined intent taxonomy.
  - **Query Reformulation Chain:** Converts user queries into structured and precise formats.

Chain-of-Thought Prompting

- Enables structured reasoning by breaking down complex tasks into sequential steps.
- **Methodology**:
  - Analyses user query contextually with object-topic pairs.
  - Leverages intent taxonomy for systematic alignment.
- **Process**:
  - Identifies main intent category from taxonomy.
  - Determines sub-intent based on specific actions or information sought.
  - Validates selections for consistency and relevance.
- **Handling Ambiguity**:
  - Prompts users for clarification if necessary.
- **Outcome**:
  - Improves precision and adaptability in intent recognition.
  - Ensures clear resolution paths for complex queries.

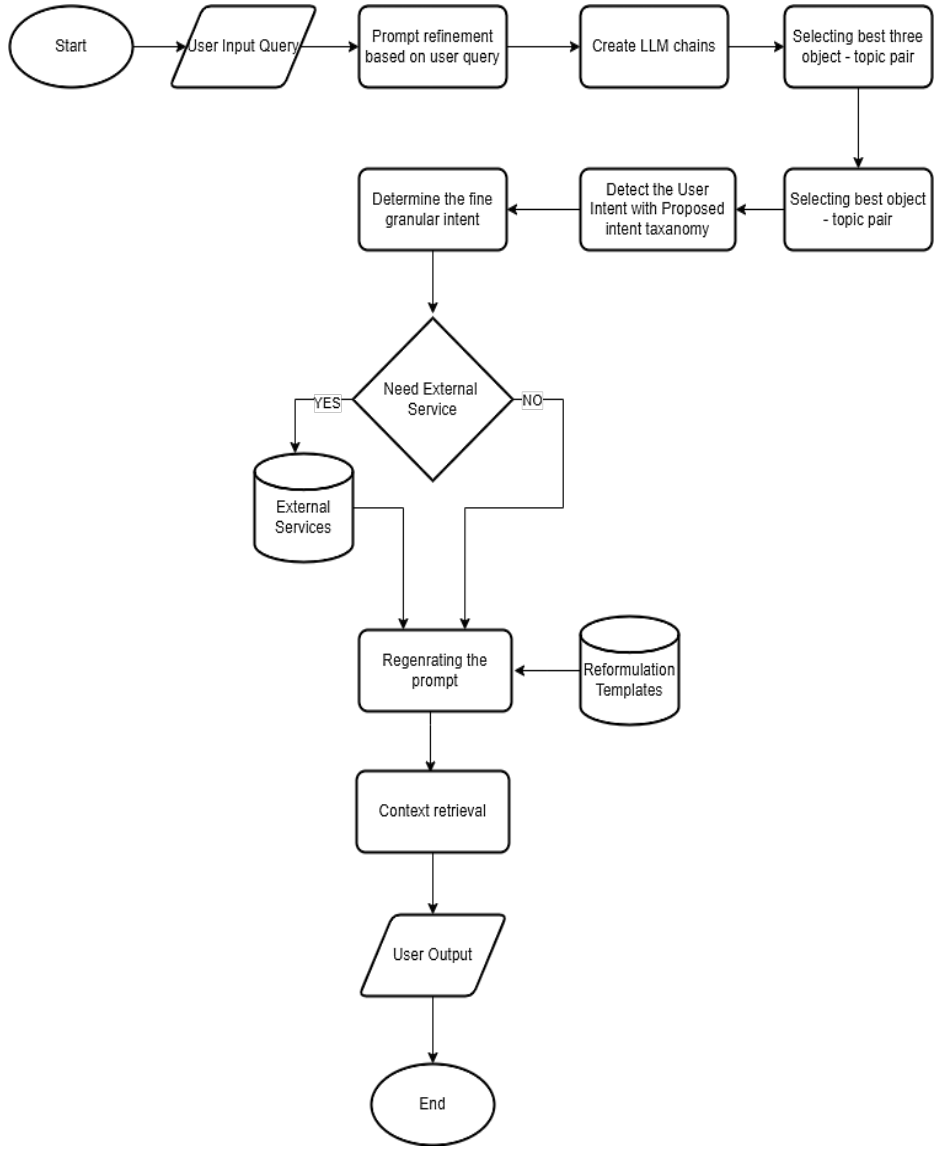
Algorithm 1 Systematic Intent Detection Process using Chain of Thought approach

Data : Best Pair P, User Query Q, Intent Taxonomy T
Result : Main Intent I_m - Sub Intent I_s
Step 1: Initial Context Analysis
Analyze the user query Q and the best pair P for explicit and implicit information
if *Q is clear and specific* **then**
  Proceed to Step 2
else
  Flag as ambiguous and proceed to Step 5
end
Step 2: Identify I_m
Refer to the T and select I_m that aligns with the user's overall goal
if *I_m accurately reflects Q* **then**
  Proceed to Step 3
else
  Reassess context or consider broader intent categories; Return to Step 2
end
Step 3: Determine I_s
Select I_s that most precisely addresses the specific action or information the user seeks
if *I_s is specific and relevant to Q* **then**
  Proceed to Step 4
else
  Reevaluate I_m or consider alternative granular intents; Return to Step 2
end
Step 4: Validate and Confirm
Re-read Q and validate that the chosen I_m and I_s align with the user's intent
if *Selected intents fully align with Q* **then**
  Finalize $I_m - I_s$ and output
else
  Reassess I_m and I_s; Return to Step 2 or Step 3
end
Step 5: Handling Ambiguity
Flag Q as ambiguous and take one of the following actions:
1. Default to a broader intent category
2. Prompt the user for more specific information

if *Ambiguity is resolved* **then**
  Return to Step 2
else
  Document uncertainty or request further clarification
end

Proposed Architecture for Intent Recognition System

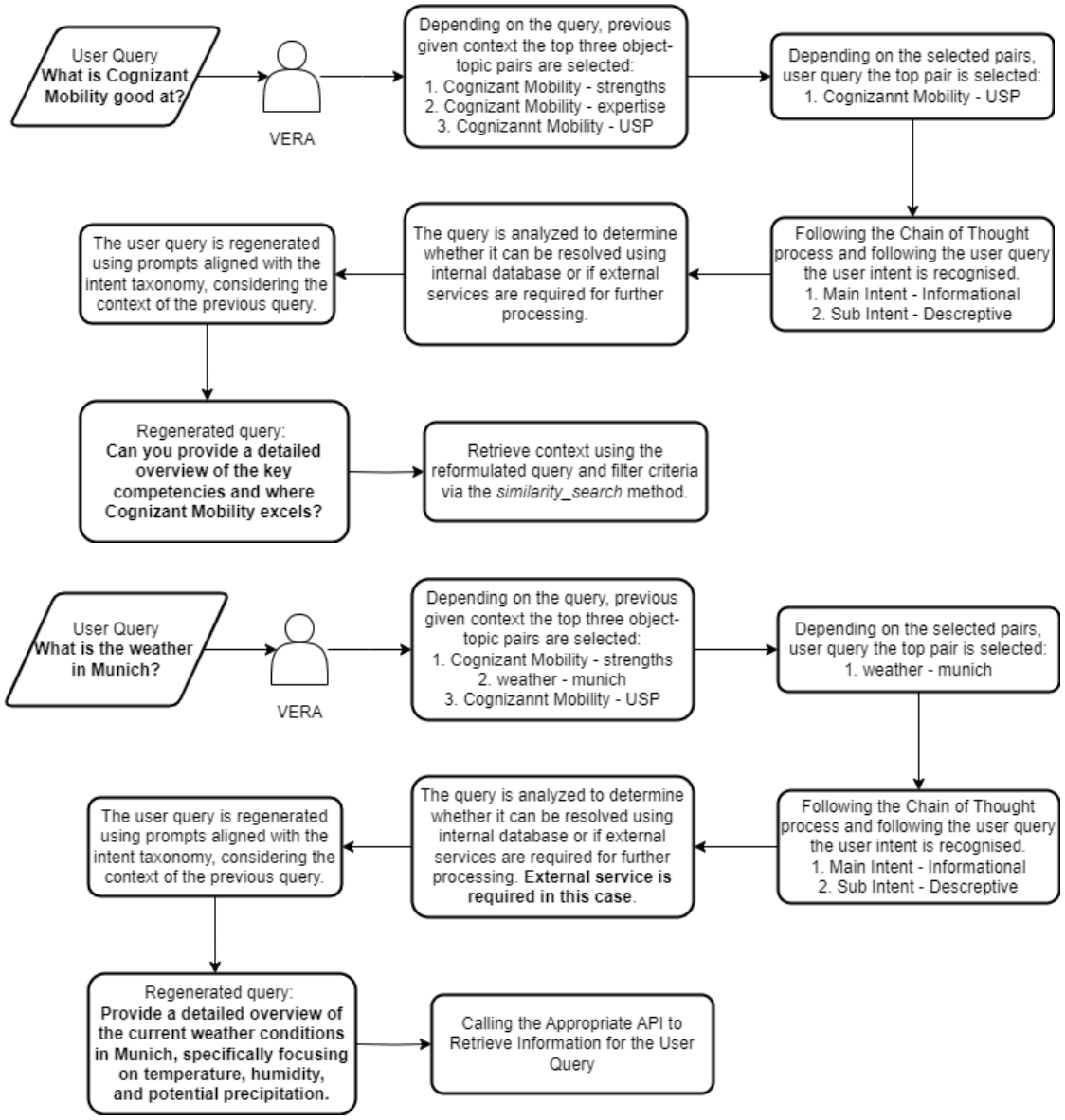
- **Foundation on Existing Framework:** Retains initial steps of receiving a user query and instantiating an LLM chain.
- **Object-Topic Pair Selection:** Identifies top three most relevant object-topic pairs, expanding possible response options.
- **Refined Pair Selection & Intent Detection:** Selects the most suitable pair and detects user intent for accurate query handling.
- **External Service Analysis:** Determines if the query requires external services (e.g., APIs) or can be resolved using the database.
- **Query Reformulation:** Reformulates queries with a template aligned to detected intent for clarity and relevance.
- **Response Generation:** Delivers contextual and user-centric responses through an enhanced query handling mechanism.



Proposed System Workflow Design

- **Retrieving Queries from Internal Database:**
  - User query processed and matched with relevant object-topic pairs in the database.
  - Chain-of-thought intent recognition identifies the intent using predefined taxonomy.
  - If no external data needed, query reformulation occurs using a template.
  - Reformulated query used for similarity search to retrieve data from the database.

- **Retrieving Queries Using External Services:**
  - System detects queries requiring external data.
  - Identified intent triggers external API calls.
  - Reformulated query sent to API, and response integrated into final user response.



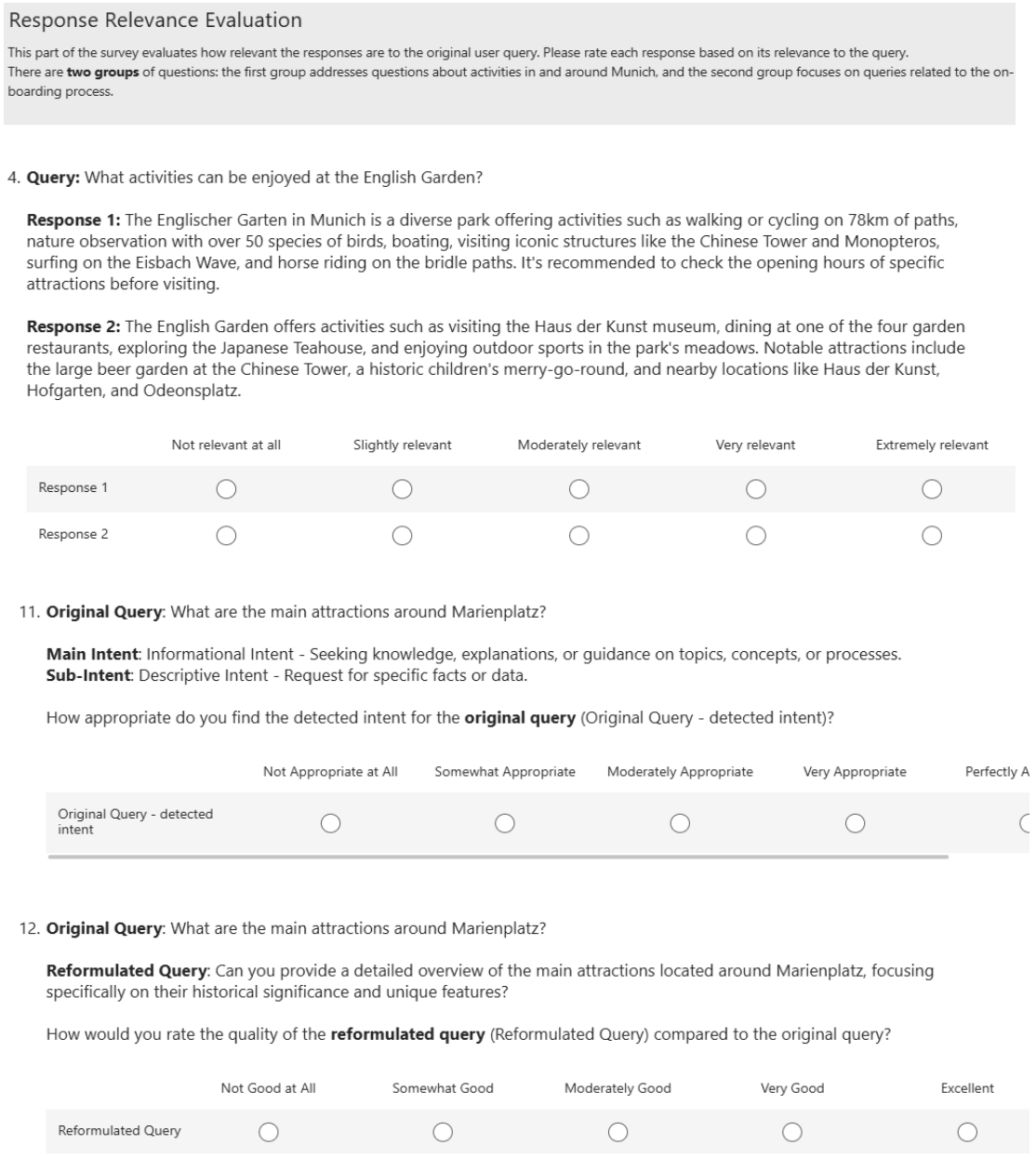
5

Evaluation Methodology

Survey Design

- **Objective of Evaluation**:
  - Assess user comprehension of intent classification.
  - Evaluate the effectiveness of the prompt reformulation technique.

- **Survey Structure**:
  - Data protection notice.
  - General chatbot usage questions.
  - Query-Response evaluation.
  - Detailed intent recognition and reformulation evaluation.
    - Intent recognition satisfaction.
    - Reformulated query assessment.
  - Overall satisfaction ratings for intent detection and reformulated query effectiveness.



Response Relevance Evaluation

This part of the survey evaluates how relevant the responses are to the original user query. Please rate each response based on its relevance to the query.
There are **two groups** of questions: the first group addresses questions about activities in and around Munich, and the second group focuses on queries related to the on-boarding process.

4. **Query:** What activities can be enjoyed at the English Garden?

Response 1: The Englischer Garten in Munich is a diverse park offering activities such as walking or cycling on 78km of paths, nature observation with over 50 species of birds, boating, visiting iconic structures like the Chinese Tower and Monopteros, surfing on the Eisbach Wave, and horse riding on the bridle paths. It's recommended to check the opening hours of specific attractions before visiting.

Response 2: The English Garden offers activities such as visiting the Haus der Kunst museum, dining at one of the four garden restaurants, exploring the Japanese Teahouse, and enjoying outdoor sports in the park's meadows. Notable attractions include the large beer garden at the Chinese Tower, a historic children's merry-go-round, and nearby locations like Haus der Kunst, Hofgarten, and Odeonsplatz.

	Not relevant at all	Slightly relevant	Moderately relevant	Very relevant	Extremely relevant
Response 1	○	○	○	○	○
Response 2	○	○	○	○	○

11. **Original Query**: What are the main attractions around Marienplatz?

Main Intent: Informational Intent - Seeking knowledge, explanations, or guidance on topics, concepts, or processes.
Sub-Intent: Descriptive Intent - Request for specific facts or data.

How appropriate do you find the detected intent for the **original query** (Original Query - detected intent)?

	Not Appropriate at All	Somewhat Appropriate	Moderately Appropriate	Very Appropriate	Perfectly A
Original Query - detected intent	○	○	○	○	○

12. **Original Query**: What are the main attractions around Marienplatz?

Reformulated Query: Can you provide a detailed overview of the main attractions located around Marienplatz, focusing specifically on their historical significance and unique features?

How would you rate the quality of the **reformulated query** (Reformulated Query) compared to the original query?

	Not Good at All	Somewhat Good	Moderately Good	Very Good	Excellent
Reformulated Query	○	○	○	○	○

Survey Criteria

- **Evaluation Approach:**
  - Compared performance of responses to **original queries** vs. **reformulated queries**.
  - Analysis conducted using **statistical methods**.

- **Paired t-Test**:
  - Primary statistical method used to compare paired results (original vs. reformulated responses).
  - Tests the null hypothesis: **Mean difference between paired results is zero**.
  - $t = \frac{X_1 - X_2}{\frac{\sqrt{s_1^2 + s_2^2 - 2\rho s_1 s_2}}{\sqrt{n}}}$

- **P-Value**:
  - Determines whether to accept or reject the null hypothesis.
  - Compared against significance level ($\alpha = 0.05$)
  - $p<0.05$: Statistically significant; evidence against null hypothesis.

6

Results

Query-Response Evaluation Results

- Evaluate the impact of reformulated queries (based on detected intents) on response quality compared to original queries.

- **Statistical Evaluation:**
  - **t-statistic:** 1.639
  - **p-value:** 0.103 (Not statistically significant, α=0.05)
  - **Findings:**
    - Reformulated queries show slightly higher mean response scores than original queries.
    - Suggests a potential improvement, though not conclusive.

Evaluation of Intent Recognition and Reformulation

- **Main Intent Categories (Out of 5):**
  - **High-performing:** Navigational (Avg: 3.75), Administrative (Avg: 3.78).
  - **Lowest-performing:** Personal Interaction (Avg: 3.32).

- **Fine-Granular Intent Categories (Out of 5):**
  - **High satisfaction:** Descriptive (Avg: 3.87), Procedural (Avg: 4.09) intents.
  - **Lower scores:** Conversational Engagement (Avg: 2.91), Reflection & Insights (Avg: 3.43).

- **Reformulated Queries Analysis (Out of 5):**
  - **Top-performing:** Procedural (Avg: 4.33), Location-Based (Avg: 4.25), and Access Management (Avg: 4.25)
  - **Lowest-performing:** Conversational Engagement (Avg: 3.13) and Consultative Intent (Avg: 3.50).

- **Statistical Summary: Intent Detection vs. Reformulation (Out of 10):**
  - **Intent Detection:** Avg: 7.54
  - **Reformulated Queries:** Avg: 8.00

7

Conclusion

Conclusion

- Enhanced intent recognition in conversational AI systems through a novel architecture combining intent taxonomies and query reformulation.

- Improved user intent detection in VERA and Onboarding AI assistant to enhance interaction quality.

- **Key Contributions**:
  - Development of a comprehensive, versatile intent taxonomy for broader applicability.
  - Integration of taxonomy with reformulation templates to deliver user-centered responses.

- Leveraged prompt engineering and reformulation templates for optimized user interaction.

Future Scope

- **Refinement of Intent Taxonomy.**
- **Advanced Query Reformulation.**
- **Broader Evaluation Frameworks.**
- **Leverage Emerging Technologies.**
- **Expand Application Domains.**
- **Multi-Language Support.**

Thank You for Your Attention!

Time for your feedback and questions ☺

Contact
Piyush Sham Patil
1634587
piyush.patil@student.uni-siegen.de
uni-siegen.de