Data specification navigator (pilot study)

1 Introduction

This report presents the results of a small-scale pilot study conducted as part of the research project for the Data Specification Navigator tool. The goal was not to establish definitive results, but to gain initial insights into efficiency, correctness, and usability.

1.1 Objective of the study

The study aims to evaluate whether the Data Specification Navigator tool enables users to efficiently create SPARQL queries that would help them get data to answer specific questions. The study compares the time it takes users to create SPARQL queries to the time it takes experts to write the queries without using the Data Specification Navigator.

2 Methodology

This chapter discusses the setup for the study.

2.1 Participants

A total of ten participants were recruited and divided into three groups.

- Group A: People who are proficient in SPARQL (4 people).
- Group B: People who do not have experience with SPARQL (5 people).
- Group C: People who do not have experience with SPARQL (1 person).

Both groups B and C have people who do not have experience with SPARQL, but their task in this study will be different.

2.2 Materials

Natural-language questions: Three questions in natural language for which corresponding SPARQL queries were required from the participants. The questions were identical for all participants. The questions were:

- List all tourist destinations that have a capacity of at least 100 people.
- List all barrier-free tourist destinations that have an elevator at the main entrance.
- List all barrier-free tourist destinations that have an elevator in the interior. For each destination, optionally list whether smoking is allowed or not.

Data specification: A file containing the description of the "tourist destination" dataset.

Software: The Data Specification Navigator tool.

2.3 Tasks

Group A: Study the data specification and write SPARQL queries for the three given questions.

Group B: Use the Data Specification Navigator tool to create SPARQL queries for the three given questions.

Group C: Learn about RDF and SPARQL, study the data specification and write SPARQL queries for the three given questions.

2.4 Procedure

Group A: This group received the data specification and the three natural-language questions. Their task was to study the given data specification and write a SPARQL query for each of the given questions.

Group B: This group received the Data Specification Navigator tool and the three natural-language questions. Participants were given the motivation behind the tool and instructions on how to use it. For each natural-language question, they were instructed to use the tool and start with the message "Show me all tourist destinations". Their task was to utilize the features of the Data Specification Navigator tool to get to the SPARQL queries that would answer the given questions.

Group C: The single participant in this group was introduced to RDF and the SPARQL language (not a deep dive, only enough information for them to complete this study). They then received the data specification and the natural-language questions. Their task was the same as group A: to study the data specification and write a SPARQL query for each of the given questions.

2.5 Data Collected

Time per question per participant and correctness of queries. In group C, the time for introduction to RDF and SPARQL was also noted.

3 Results

This chapter contains the data collected (time and correctness) and a few qualitative feedback from the participants of the study. The terms "question 1", "question 2" and "question 3" will sometimes be abbreaviated to Q1, Q2 and Q3.

3.1 Time per question

Table 1 and Chart 1 show the average times per group for each question (table 1 shows minutes and seconds while chart 1 shows times in seconds). There was only one participant in group C so the values reflect that individual's times. For groups A and C, the time for question 3 was faster than for question 2, whereas for group B the time needed for question 2 was shorter than the time needed for question 3.

Chart 2 shows the time distribution per group for question 1. Chart 3 and Chart 4 show the time distribution per group for questions 2 and 3 respectively. Group C consisted of a single participant so the boxplots show a single value without variability.

Table 1: Average time for each question.

Question	Group A (experts)	Group B (tool users)	Group C (novice learner)
Q1	2m20s	6m52s	5m11s
Q2	2m43s	8m21s	10m43s
Q3	2m11s	11m01s	6m47s

Average time (in seconds) per question

	Question 1	Question 2	Question 3	
Group A	140	163	131	
Group B	412	501	661	
Group C	311	643	407	

Group A ● Group B ● Group C ●

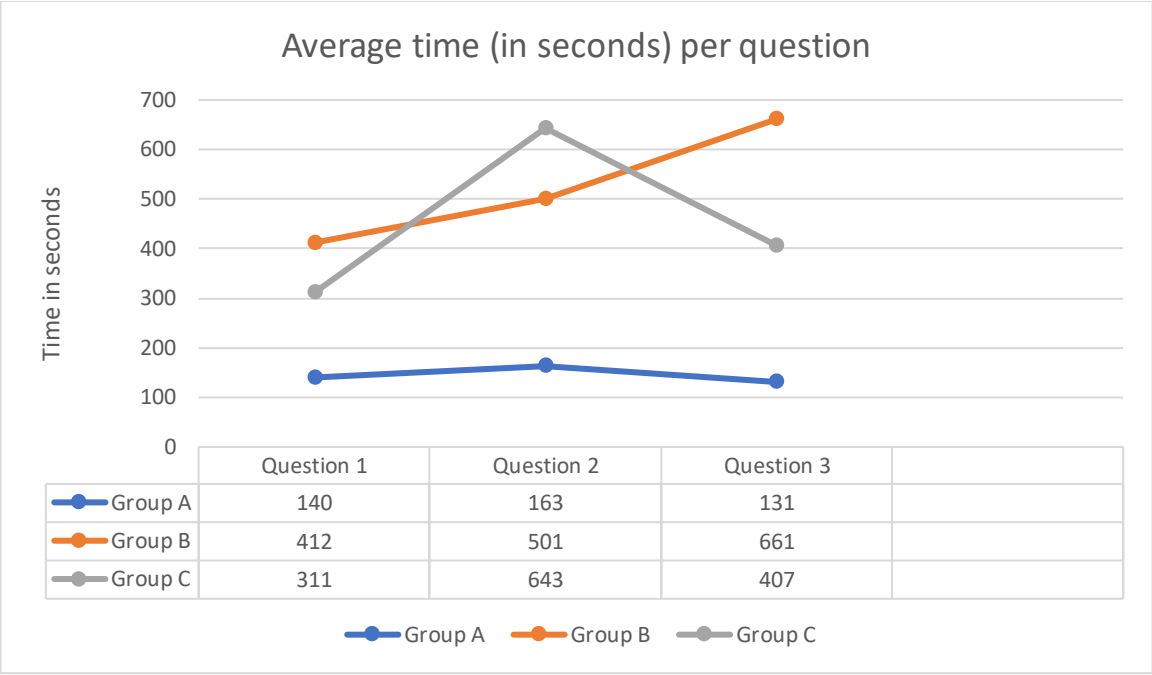## Time distribution for question 1
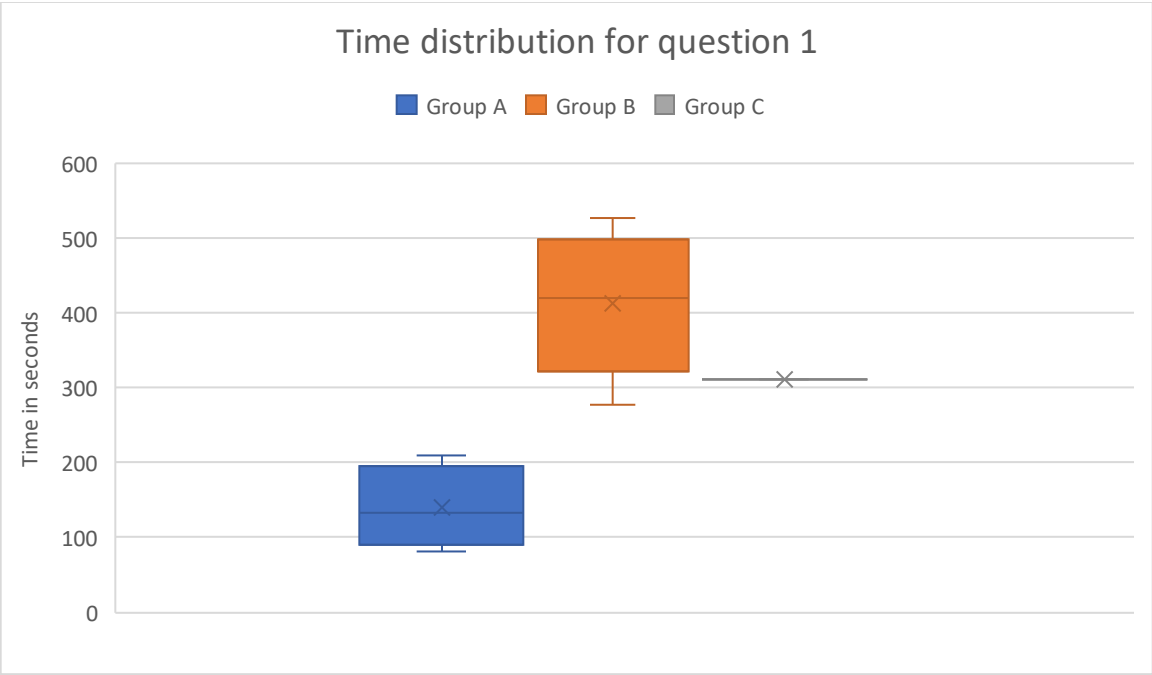
Group A ■ Group B ■ Group C ■

Chart 3: Time distribution for question 2 per group. Group C consisted of only one participant; therefore, the boxplot shows a single value without variability.
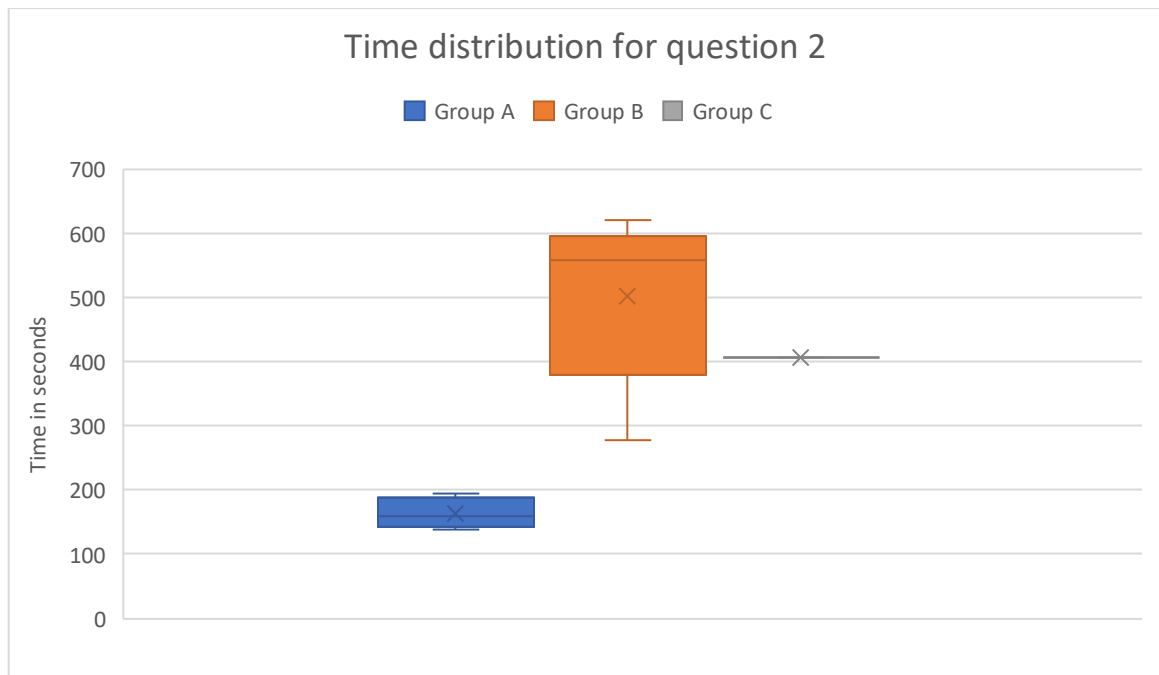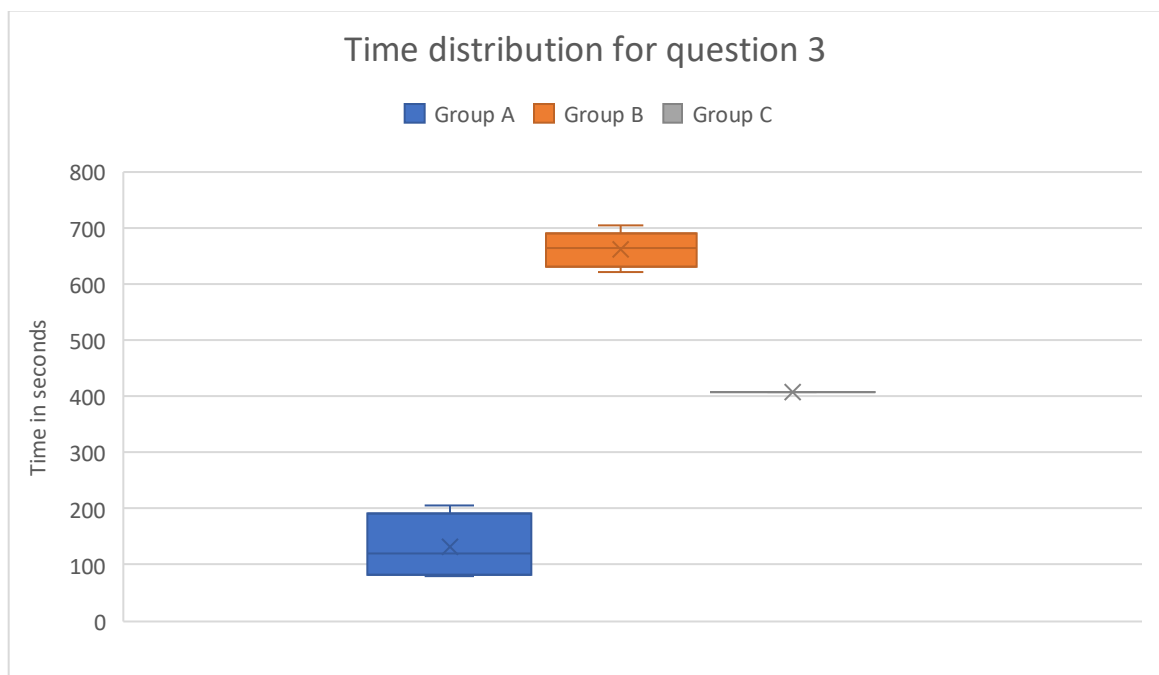
Chart 4: Time distribution for question 3 per group. Group C consisted of only one participant; therefore, the boxplot shows a single value without variability.

3.2 Correctness of queries

Group A (experts): All queries correct.

Group B (tool users): Q1 and Q2 correct for all participants. Q3 incorrect for all participants.

Group C (novice learner with brief training): All queries correct.

3.3 Training time for group C

The single participant in group C completed a brief training on RDF and mainly SPARQL. The session consisted of three parts:

- Introduction to RDF and OWL: 10 minutes.
- Introduction to SPARQL: 10 minutes.
- Hands-on SPARQL practice (tiny examples): 10 minutes.

Overall, the session took about 30 minutes.

3.4 Qualitative feedback

One participant from group B noted that the color of the UI is not suitable for color-blind people.

Group B participants shared the sentiment that the SPARQL query in the chatbot's answer is taking too much space when they do not understand it because they are not familiar with SPARQL. It would be better to move it elsewhere or add a "show/hide SPARQL" button. They have additionally noted that the data specification exploration part of the tool is interesting.

One participant in group A noted that they did not have to study the whole data specification. They simply looked up "elevator" in the data specification and used what they needed.

4 Discussion

This chapter discusses the measured times and offers some interpretations of the results.

4.1 Efficiency and correctness

Experts (group A) outperformed both other groups in speed and correctness. The tool (Group B) enabled non-experts to produce correct queries for Q1 and Q2 but failed in Q3. The single novice learner in Group C achieved correctness on all queries and their time was faster on Q1 and Q3 but slower in Q2 compared to group B's average.

All participants in group B had the same mistakes in the query for question 3:

- Missing property "elevator in interior".
- Missing the OPTIONAL modifier for "smoking is allowed".

The missing "elevator in interior" is caused by the tool not suggesting this property to the users. Users dealt with this in various ways. One user tried starting from the beginning by sending the "show me all tourist destinations" message. Others tried resending their current message. Ultimately, all of them gave up thinking the "elevator in interior" is not in the data.

The missing OPTIONAL modifier is likely caused by users not knowing how exactly the "Add as OPTIONAL" toggle in the UI works. This indicates a need to make the toggle's function clearer to users.

4.2 Comparison between Q2 and Q3

An interesting pattern emerged in the results: participants in groups A and C typically solved Q3 faster or just as fast as Q2. This can be explained by the structural similarity of the two questions. Q2 and Q3 both required joining classes "tourist destination" and "barrier-free access". The only difference in Q3 was that it required an additional OPTIONAL modifier for the "smoking is allowed" attribute. Once participants had figured out the solution strategy for Q2, they were able to adapt it quickly for Q3, resulting in shorter times.

For Group B, however, the Data Specification Navigator interface demanded a similar number of interaction steps for both Q2 and Q3. Even if participants recognized the similarity between the two questions, the tool did not reduce the effort needed to get to the final answer. This coupled with the previously mentioned issue of the tool not giving a relevant suggestion resulted in longer time spent of Q3 for group B.

4.3 Educational approach

The group C participant demonstrated that with roughly 30 minutes of introduction, a novice can successfully produce correct queries for simple questions. It is possible that training users in SPARQL yields higher accuracy but demands time investment. However, the evidence is inconclusive as the group contained only a single participant.

5 Limitations

The design of the user study had several limitations.

- **Number of participants**: The study involved only 10 respondents, which limits the generalizability of the results. Notably, group C consisted of only a single participant.
- **Number of tasks**: Only three simple questions were used as tasks. Repeating the study with more questions and more complex questions might yield more interesting results.
- **Focus of evaluation**: The study setup focused on having the tool users (group B) start with the question "Show me all tourist destinations" and then build towards a predefined question. This does not fully reflect the main purpose of the software, which is to support *data specification exploration*. In real use cases, users would typically come with a question, and the tool would help them discover relevant classes and properties by suggesting possible extensions. By constraining the tasks to predefined targets, the study may have underestimated the exploratory strengths of the tool.

6 Future work

Several directions remain for future studies.

- **Larger-scale evaluation**: Increasing the number of participants, particularly in group C, would provide more robust insights into how novice training compares to tool usage.
- **Task realism**: The current study required users to reach predefined queries. Future experiments should allow more open-ended exploration to better reflect the actual purpose of the tool: assisting users in *discovering and exploring* data specifications.
- **Usability testing**: Collecting systematic usability feedback (e.g., via questionnaires or open-ended interviews) could help identify UI issues and improve user experience.

7 Conclusion

This pilot study compared three groups: SPARQL experts (group A), non-experts using a tool (group B), and a novice with a short training session (group C).

The findings show that:

- Experts remain the most efficient and accurate in constructing SPARQL queries.
- The tool allowed non-experts to produce correct queries for simple and moderately complex tasks, though it failed on a more complex query involving an OPTIONAL modifier.
- A novice learner, after roughly 30 minutes of training, was able to construct all queries correctly. They spent more time on a moderately complex question but once they had mastered the strategy, they were able to solve structurally similar questions faster than tool users.

In conclusion, while experts remain unmatched in efficiency, the study showed the potential of the Data Specification Navigator tool to help retrieve data for users' questions.