

University of Tartu
Faculty of Social Sciences
Narva College
Information Systems Development

Ivan Puhachov

**USER BEHAVIOR AND AI USAGE ANALYSIS
IN A FITNESS APPLICATION
USING TELEMETRY DATA**

Final Thesis

Supervisor: Andre Säask, M.Sc.

Narva 2026

Olen koostanud töö iseseisvalt. Kõik töö koostamisel kasutatud teiste autorite tööd, põhimõtetelised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on viidatud.

.....

Töö autori allkiri ja kuupäev

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Ivan Puhachov,

(autori nimi)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose **Kasutajakäitumise ja tehisintellekti kasutamise analüüs treeningrakenduses telemeetriaandmete põhjal**

(lõputöö pealkiri)

mille juhendaja on Andre Säask,

(juhendaja nimi)

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Ivan Puhachov

11.05.2026

ABSTRACT

This thesis focuses on the analysis of user behavior and AI usage in a fitness application using telemetry data. Modern applications increasingly use AI features, but there is limited understanding of how users interact with them in real-world conditions.

The aim of this study was to collect and analyze telemetry data, including AI request frequency, response time, user activity patterns, and basic demographic information. The system was implemented using a React frontend, a FastAPI backend, and a PostgreSQL database.

The analysis was based on application data collected during system usage from 18 March 2026 to 11 May 2026. During this period, the system recorded 14 users, 60 sessions, 910 telemetry events, and 164 AI requests. The results showed that AI usage differed between user groups, with the highest request activity observed in the 18–24 age group. The average AI response time was 3424.0 milliseconds.

The results demonstrate that telemetry data can be used to identify preliminary usage patterns, evaluate AI system performance, and compare user activity across demographic segments. Due to the small dataset size, the findings remain descriptive and should not be interpreted as statistically validated relationships.

TABLE OF CONTENTS

ABSTRACT	4
TERMS AND ABBREVIATIONS.....	7
INTRODUCTION	9
1. THEORETICAL BACKGROUND.....	11
1.1 Conversational Artificial Intelligence in Consumer Applications	11
1.2 User Interaction with AI Systems and Behavioral Data	11
1.3 Telemetry Data and Analytical Approaches in AI Systems.....	12
2. METHODOLOGY.....	14
2.1 Research Design.....	14
2.2 System and Data Collection Design.....	15
2.3 System Architecture Implementation.....	16
2.4 Data Storage and Research Variables.....	19
2.5 Class Diagram	21
2.6 Analytical Procedure.....	21
2.7 Ethics and Data Protection	22
2.8 Personal Contribution	23
3. RESULTS AND ANALYSIS.....	25
3.1 Overview of Collected Data	25
3.2 AI Usage Patterns	26
3.3 System Performance Analysis.....	26
3.4 Demographic-Based Analysis.....	27
4. DISCUSSION.....	29

4.1 Interpretation of Results	29
4.2 Implications for System Design	29
4.3 Limitations	30
4.4 Future Work	30
CONCLUSION	32
KOKKUVÕTE	34
REFERENCES	36

TERMS AND ABBREVIATIONS

AI (Artificial Intelligence) – a field of computer science focused on creating systems capable of performing tasks that typically require human intelligence, such as decision-making, pattern recognition, and natural language understanding.

Conversational AI – a type of artificial intelligence that enables systems to interact with users through natural language using techniques such as large language models and chat-based interfaces.

Telemetry – automatically collected data that records user actions and system events during application usage, including interactions, requests, and responses.

Token – a unit of text processed by language models, used to measure input and output size in AI interactions.

Response Time – the time required for the AI system to generate a response after receiving a user request, typically measured in milliseconds.

Descriptive Analytics – a type of data analysis that summarizes historical data using metrics such as counts, averages, and distributions to identify patterns in user behavior.

User Segmentation – the process of grouping users based on shared characteristics, such as demographics or behavior, to identify different usage patterns.

API (Application Programming Interface) – a set of rules and protocols that allow different software components to communicate with each other.

JWT (JSON Web Token) – a token-based authentication mechanism used to securely identify users in the system.

PostgreSQL – a relational database management system used to store application data, including telemetry and AI interaction metadata.

JSONB – a binary JSON data format in PostgreSQL used to store flexible and structured event data.

OpenAI API – an external AI service used by the system to process conversational requests and generate responses.

INTRODUCTION

Artificial intelligence (AI) is widely used in modern applications, including fitness apps. These systems provide recommendations, answer questions, and support user decisions. However, there is still limited understanding of how users actually interact with AI features in real applications.

Many existing studies are based on surveys or controlled experiments, which do not always reflect real user behavior. At the same time, modern applications generate telemetry data that records user actions, system responses, and usage patterns. This data creates an opportunity to study real interactions at scale.

Fitness applications are a suitable context for this type of analysis because they combine AI features with user-specific data such as age, gender, and physical characteristics. These factors may influence how often users interact with AI and how they use it. However, the relationship between these factors and AI usage is still not well studied using real application data.

This thesis focuses on analyzing user interaction with AI in a fitness application using telemetry data. The goal of the study is to identify usage patterns and examine how they are related to user characteristics and engagement.

The objectives of this thesis are:

- To design a system for collecting and storing telemetry data related to AI interactions.
- To analyze key usage metrics such as request frequency, response time, and activity patterns.
- To examine how demographic characteristics are associated with AI usage behavior.
- To identify differences between user groups based on usage patterns.
- To propose practical recommendations for improving AI features based on data.

The study is guided by the following research questions:

- What patterns can be observed in user interaction with AI features?
- How frequently do users interact with AI, and how does this vary between user groups?

- How are demographic factors associated with AI usage?
- What is the relationship between user engagement and AI interaction?
- How can telemetry data be used to improve AI-driven features?

The structure of this thesis is as follows: chapter 1 presents the theoretical background, chapter 2 describes the methodology, chapter 3 presents the analysis, followed by discussion and conclusions.

1. THEORETICAL BACKGROUND

1.1 Conversational Artificial Intelligence in Consumer Applications

Conversational artificial intelligence is widely used in modern consumer applications. Instead of static interfaces, many systems now allow users to interact through natural language using chat-based AI. These systems can answer questions, provide recommendations, and support decision-making.

Conversational AI refers to technologies that enable software systems to communicate with users in a natural conversational way. These systems are designed to simulate human interaction through text or speech, making communication more intuitive and flexible¹.

In practice, conversational AI systems are often based on large language models. They are used in various applications to process user input, generate responses, and adapt to user needs.

1.2 User Interaction with AI Systems and Behavioral Data

Understanding how users interact with AI systems is an important research topic, especially in real-world applications. Many traditional studies rely on surveys or controlled experiments, which may not fully reflect actual user behavior. In contrast, modern applications generate behavioral data that captures real user actions in real time.

AI-based systems enable continuous interaction between users and applications. This makes it possible to collect detailed data about user behavior, such as request frequency, response time, session duration, and time-of-day usage patterns. These metrics help to understand how users engage with AI features and how these systems are used in practice.

Conversational agents are widely used across different domains and provide valuable insights into user needs and behavior. By analyzing interaction data, it becomes possible to improve user experience and optimize system performance².

¹ Mariani, M. M., Hashemi, N., & Wirtz, J. (2023).

² Mariani, M. M., Hashemi, N., & Wirtz, J. (2023).

Telemetry data plays a central role in this process. It refers to automatically collected data that records user actions and system events, including requests, responses, and application interactions. This data is typically stored in real time and provides a detailed view of how users interact with digital systems.

This approach is consistent with the concept of digital behavioral data collection, where user behavior is continuously recorded through system logs and device-based data sources.³

In this study, telemetry data is used to analyze relationships between user characteristics and AI usage. By combining behavioral data with demographic attributes such as age, gender, and physical parameters, it becomes possible to identify patterns, differences between user groups, and factors influencing engagement.

1.3 Telemetry Data and Analytical Approaches in AI Systems

Telemetry data in AI applications is used not only for monitoring system performance but also for analyzing user behavior and supporting data-driven decisions. In modern systems, telemetry is typically stored in structured formats such as event logs or databases, which allows efficient querying and analysis⁴

One of the main approaches used in this context is *descriptive analytics*. This method focuses on summarizing user behavior using metrics such as request frequency, response time, session duration, and usage intensity. These metrics provide a clear overview of how users interact with the system⁵.

Another commonly used approach is the analysis of relationships between variables in a descriptive sense. This includes comparing user groups and examining how differences in demographic characteristics may be associated with variations in AI usage.⁶

³ Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016).

⁴ Microsoft. Application Insights telemetry data model (2026).

⁵ Qlik. What is descriptive analytics?

⁶ Chew S, Abd Razak SS, Lo YT. The relationship between artificial intelligence usage and academic performance of students at a Malaysian public university. (2025).

User segmentation is also an important analytical approach. It involves grouping users based on shared characteristics, such as demographics or behavior patterns. This allows the identification of different usage patterns and supports better system adaptation to user needs.⁷

In addition, *data visualization* plays an important role in analysis. Charts and dashboards help present data in a clear and understandable way, making it easier to identify trends and differences between user groups⁸.

In this study, these approaches are used to analyze how users interact with AI features in a fitness application. The combination of telemetry data, statistical analysis, and visualization provides a structured way to understand user behavior and improve AI-driven systems.

⁷ Manero A, et al. Profile segmentation: Clustering approach based on demographics and behavioral traits. *SAGE Journals*. 2025.

⁸ Lawson-Body A. Anomaly identification through data visualization. *Issues in Information Systems*. 2023;24(4):161-175.

2. METHODOLOGY

2.1 Research Design

This study is designed as an applied, data-driven investigation of user interaction with AI features in an implemented fitness application. The research follows a quantitative and exploratory approach, focusing on the analysis of behavioral data generated during actual system usage rather than controlled experimental conditions.

The design is grounded in the implemented GymBRO system, which enables automatic collection of user interaction data during normal application use.

The object of analysis in this study is user interaction with the GymBRO application, specifically the way users engage with its conversational AI functionality and other application features. Instead of relying on self-reported data or experimental tasks, the study is based on digital traces produced during system usage. These traces include behavioral telemetry events, AI request metadata, usage timing, and user profile attributes that enable demographic segmentation.

The research design can be characterized as observational, as it does not involve manipulation of variables or assignment of users to experimental groups. All data is collected passively as part of the system's operation. The dataset consists of automatically recorded application events, structured AI interaction metadata, and user profile information stored in the system database.

The analytical approach is based on descriptive and comparative methods such as aggregation, segmentation, and comparison of usage metrics. These methods are used to examine patterns of AI usage, differences between user groups, and general trends in user engagement and system performance.

This design is appropriate for the objectives of the thesis, as it allows the study of user behavior during normal application use. However, the design does not include qualitative data collection, experimental manipulation, or analysis of full conversational content, and therefore the study is limited to behavioral and operational aspects of AI usage.

2.2 System and Data Collection Design

The GymBRO system is designed so that research-relevant data is collected during normal application use rather than through a separate research interface.

On the client side, the application records user actions as behavioral telemetry events. These events are generated when users interact with the system, for example during authentication, onboarding, navigation between screens, use of the chat interface, AI request initiation and completion, and interaction with features such as meals, training, and sleep tracking. These events are sent asynchronously to the backend through authenticated HTTP requests to the `/analytics/events` endpoint.

Frontend telemetry is sent in a best-effort manner and is not strictly guaranteed to be delivered in all cases, as event delivery is asynchronous and does not block user interaction.

The system distinguishes between two types of collected data. The first type is frontend-generated behavioral telemetry, which represents user interaction with the application interface. Stored telemetry records include an event type, timestamp, authenticated user identifier, and additional contextual information stored in a semi-structured JSONB payload.

The second type is backend-generated operational and AI interaction data, which is produced during request processing. Each AI request is processed through a backend pipeline where the system performs authentication, manages session state, applies usage limits, and communicates with the external AI service.

During this process, structured metadata is recorded for successful AI requests, including timestamps, token usage (prompt, completion, and total tokens), response time, and model version. This metadata is stored in the database as part of the AI interaction records.

In addition to successful AI requests, the system records backend-generated operational events, such as rate-limit violations and AI provider failures. These events are logged directly within the AI request processing pipeline using the shared backend event logging function.

The overall data collection process therefore consists of two complementary streams: a behavioral telemetry stream capturing user actions, and an AI interaction stream capturing technical characteristics of AI usage and system behavior.

Backend-generated AI interaction metadata is recorded synchronously during request processing, ensuring consistent recording of successful AI request data. In contrast, frontend telemetry is asynchronous and best-effort.

This design enables the collection of structured behavioral and system-level data suitable for quantitative analysis.

2.3 System Architecture Implementation

The GymBRO system was implemented as a full-stack application consisting of a React-based frontend, a FastAPI backend, and a PostgreSQL database.

The frontend application is responsible for user interaction and context preparation. When a user sends a message to the AI assistant, the request is constructed in the chat interface by combining the user input, previous conversation history, and a dynamically generated system prompt. This combined message context is then sent to the backend via an authenticated HTTP request to the `/ai/query` endpoint.

The processing of AI requests is implemented in the backend route `/ai/query`, which orchestrates authentication, session management, rate limiting, and communication with the external AI service.

Each request follows a defined pipeline. First, the user is authenticated using a token-based mechanism. Next, the system applies rate limiting using a dedicated quota tracking table. After that, a session record is created or validated to group requests into sessions and validate session ownership; the conversational message context itself is assembled on the frontend and sent with each request. The request is then forwarded to the external AI service via the OpenAI Chat Completions API.

During processing, the system measures response time and extracts token usage statistics. After receiving the AI response, the system stores structured metadata about the interaction in the `ai_requests` table. The backend PostgreSQL database stores only structured metadata about AI interactions and does not persist full prompt or response text in the `ai_requests` table.

In addition to AI request handling, the system implements a telemetry mechanism for tracking user behavior. On the frontend, user actions trigger telemetry events, which are sent asynchronously to the backend via the `/analytics/events` endpoint.

The backend processes these events and stores them in the `telemetry_events` table. In contrast, backend-generated operational events, such as rate-limit violations and provider failures, are recorded directly within the AI request pipeline using a dedicated logging function, without passing through the telemetry ingestion endpoint.

The database schema is designed to support both relational and event-based data analysis. The `users` table stores demographic attributes, which are linked to both AI interaction data and telemetry events. The `ai_sessions` table represents session grouping of AI requests, while the `ai_requests` table contains structured metadata about AI usage. The `telemetry_events` table serves as a general event log capturing user behavior across the application.

This architecture enables a unified analytical dataset where user characteristics, behavioral events, and AI interaction metrics are connected through shared identifiers.

The architecture of the system is illustrated in Figure 1.

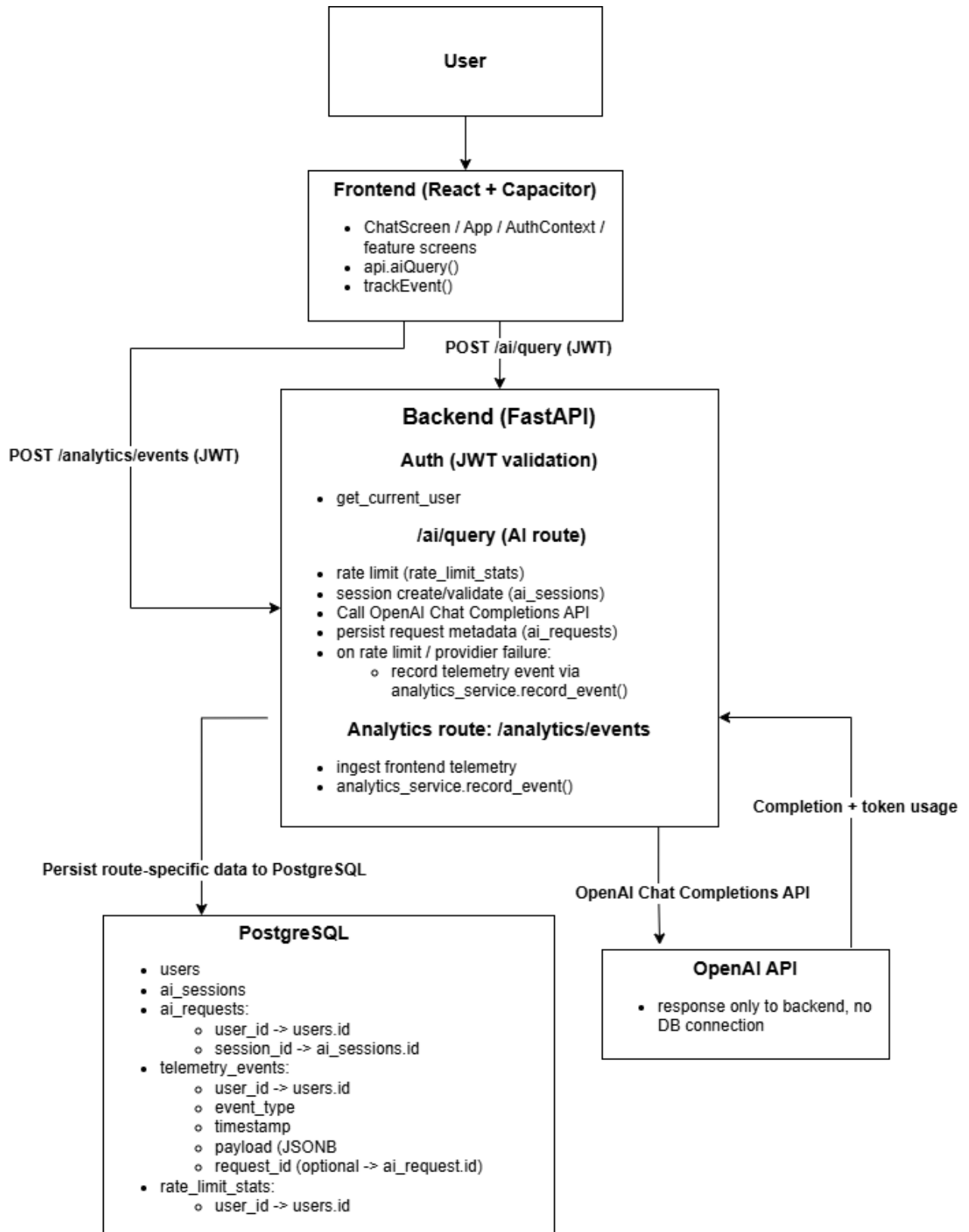


Figure 1. AI and Telemetry Data Flow Architecture of the GymBRO system

2.4 Data Storage and Research Variables

The database structure relevant to analytics and user segmentation is presented in Figure 2.

The data used in this study are stored in a PostgreSQL database and consist of both structured relational records and semi-structured event-based data. The main sources relevant for analysis include user profile data, behavioral telemetry events, AI interaction metadata, session information, and rate-limiting data.

From a research perspective, the dataset combines demographic attributes, behavioral telemetry, and AI interaction metadata into a unified analytical structure.

Demographic variables are stored at the user level and include attributes such as age group, gender, weight, and height. The age group variable is derived from the raw age attribute during profile processing.

AI usage variables are derived from AI request and session records. These include timestamps, session identifiers, model version, token usage, and response time. Rate-limiting data reflects operational quota usage and is used to analyze constraints on AI interaction frequency.

Behavioral variables are captured through telemetry events, which include event type, timestamp, user identifier, and contextual information stored in a flexible JSONB payload.

A key characteristic of the dataset is the way these variables are linked. User identifiers connect demographic data with both telemetry events and AI interaction records, while request identifiers allow certain telemetry events to be associated with specific AI requests.

This structure enables analysis at multiple levels, including individual events, AI requests, sessions, and overall user behavior. The central backend database stores AI interaction metadata rather than full conversational transcripts, which limits the thesis analysis to usage patterns and system performance.

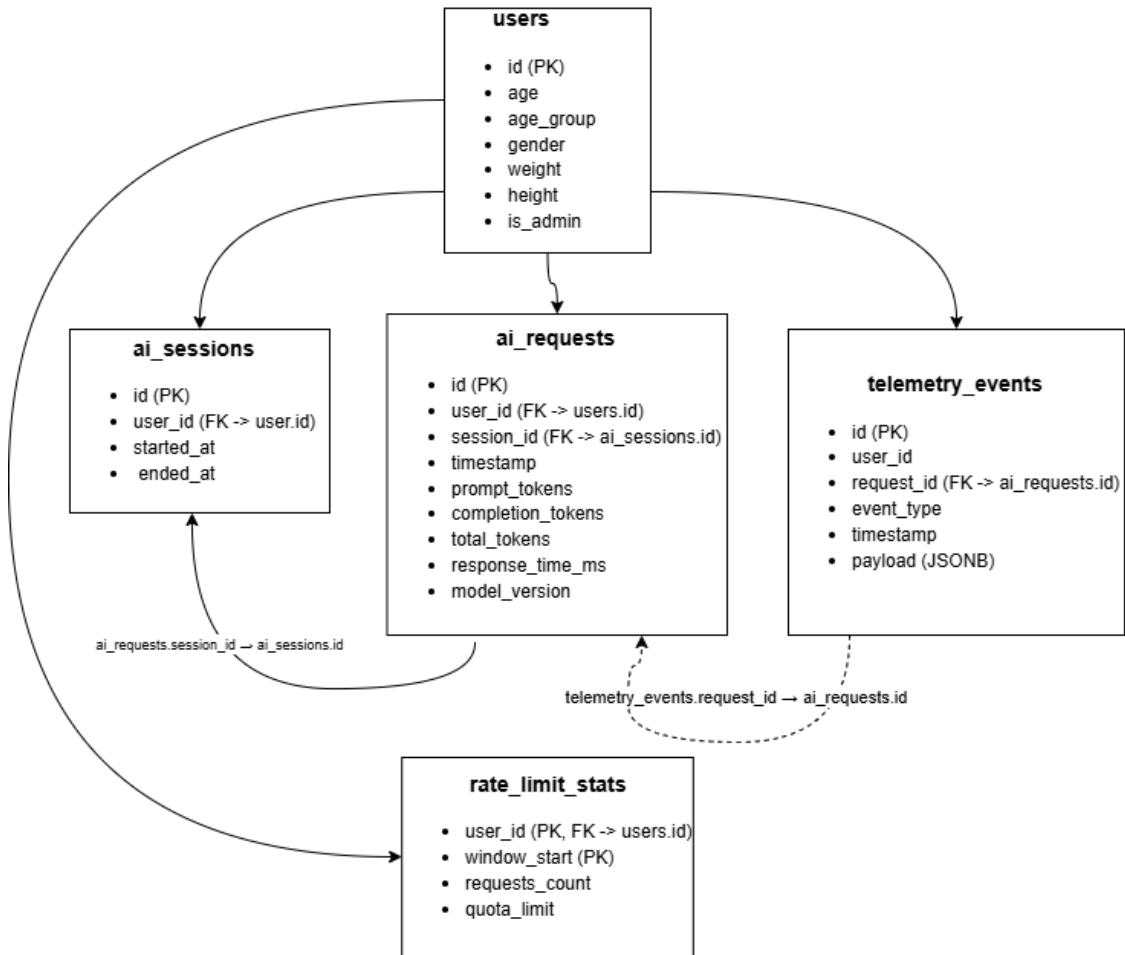


Figure 2. Database schema for analytics and user segmentation in the GymBRO system

Figure 2 presents the core database entities used in the study. Demographic attributes used for segmentation are stored in the users table. AI interaction metadata is stored in ai_requests, behavioral event logs are stored in telemetry_events, session grouping is represented by ai_sessions, and quota-related operational data is stored in rate_limit_stats.

2.5 Class Diagram

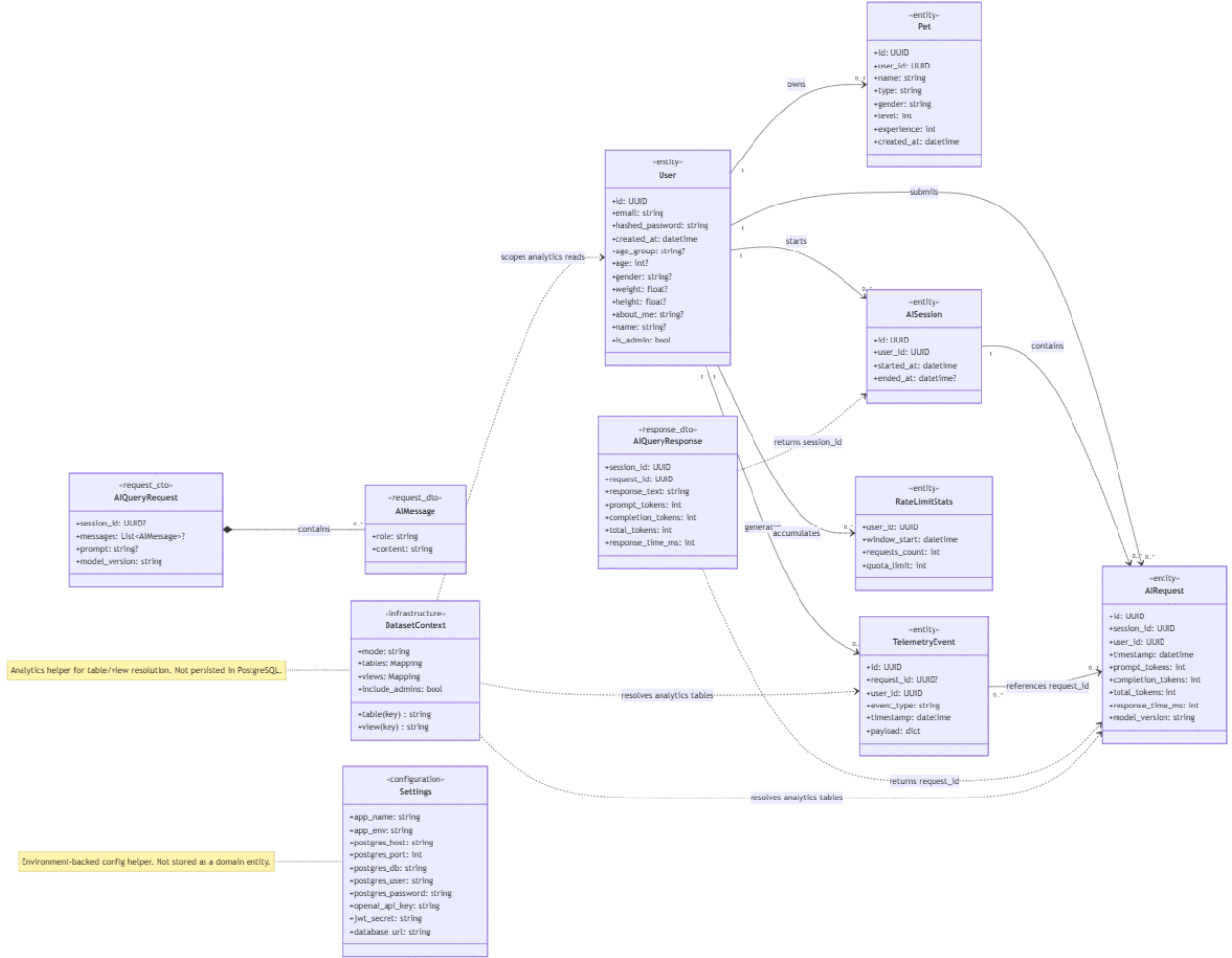


Figure 3. Backend class diagram of the GymBRO system

The class diagram shows the main backend domain entities and API data models used in the GymBRO system, including users, pets, AI sessions, AI requests, telemetry events, rate-limiting statistics, and the request and response models used for AI and analytics operations.

2.6 Analytical Procedure

The analytical procedure in this study was based on descriptive and comparative analysis of telemetry and AI interaction data collected from the GymBRO application.

The initial stage focuses on summarizing system usage through counts, totals, and averages. This includes metrics such as the number of AI requests, the number of active AI users, average requests per user, average token usage, and average response time.

The analysis then incorporates time-based aggregation, grouping data by intervals such as hour, day, week, or month.

Demographic segmentation is performed using age group, gender, weight, and height. For each segment, metrics such as request frequency, token usage, and response time are calculated.

Comparison-based analysis is used to identify differences between user groups. This includes comparing the most and least active segments. The analysis remains descriptive and does not include formal statistical inference or correlation testing.

Event-based telemetry data is used to support the overall analysis of application activity and to verify that user interactions are recorded by the system. Detailed feature-level analysis can be extended in future work by grouping telemetry events according to specific application functions.

Performance analysis is conducted using AI interaction metadata, including response time and token usage. The system also supports reliability analysis through backend-generated operational events, such as provider failures and rate-limit incidents, although detailed error-type analysis is outside the main scope of the final empirical results. The results are presented through a dashboard-based visualization approach, enabling clear interpretation of trends, distributions, and segment comparisons.

Overall, the analytical procedure is aligned with the structure of the collected data and the capabilities implemented in the system.

2.7 Ethics and Data Protection

This study is based on data collected automatically during the use of the GymBRO application. The data includes telemetry events, AI interaction metadata, and user profile attributes that are required for system functionality. The study does not involve direct interaction with participants, surveys, or manual data collection.

From an ethical perspective, the thesis presents only aggregated and de-identified results. Individual users are not identified in the results, and no personally identifiable information is presented in the thesis. However, the raw application database contains user-related

records linked to user identifiers, which are necessary for normal system operation and analytical processing.

The system design supports data minimization in the backend PostgreSQL database. The backend PostgreSQL database stores AI interaction metadata, such as token usage, response time, timestamp, and model version, but does not persist full conversational transcripts in the `ai_requests` table. This limits the scope of the thesis analysis to technical and behavioral aspects of AI usage.

Data collection is performed as part of normal application operation and is necessary for system functionality and analytics. Within the scope of this thesis, the collected data is used only for research and analytical purposes. The results are presented in aggregated form to reduce the risk of identifying individual users.

Overall, the study follows basic principles of data protection, including minimization of stored AI conversation content in the backend database, avoidance of personal identification in the thesis results, and responsible use of application-generated data for research purposes.

2.8 Personal Contribution

This thesis is based on the development of the GymBRO application, which was implemented as a collaborative project. The author of this thesis was primarily responsible for the design and implementation of the backend system, the AI interaction pipeline, and the analytics subsystem.

The main contributions include the development of the FastAPI backend, design of the PostgreSQL database schema, implementation of telemetry data collection and storage mechanisms, and creation of analytical endpoints and SQL views for processing user interaction data.

The author also implemented the AI request handling logic, including session management, rate limiting, integration with the OpenAI API, and storage of AI interaction metadata such as token usage and response time.

The frontend application and user interface were developed collaboratively, with major frontend components implemented by another contributor. Therefore, this thesis focuses

specifically on the backend architecture, analytics logic, telemetry infrastructure, and data-processing components developed by the author.

The contribution of the author can be verified through the version control history of the project repositories. Links to the repositories are given in the references part of this thesis.

3. RESULTS AND ANALYSIS

This chapter presents the results of the analysis of telemetry and AI interaction data collected from the GymBRO application. The analysis is based on data collected during the period from 18 March 2026 to 11 May 2026. The dataset includes application usage records, AI request metadata, user demographic attributes, sessions, and telemetry events.

The results are descriptive in nature. Due to the relatively small number of users, the analysis does not aim to establish statistically significant relationships. Instead, it focuses on identifying observable patterns in AI usage, user activity, demographic differences, and system performance.

3.1 Overview of Collected Data

During the selected analysis period from 18 March 2026 to 11 May 2026, the GymBRO system collected data from 14 users. The dataset included 60 sessions, 910 telemetry events, and 164 AI requests.

Out of 14 registered users, 13 interacted with the AI functionality during the selected period. The average number of AI requests per active user was 12.6. This indicates that the AI feature was used by most users in the dataset and that active users generated multiple requests on average.

The average token usage per AI request was 3078.4 tokens, while the average response time was 3424.0 milliseconds. The relatively high token usage should be interpreted in the context of the system design. Each AI request includes not only the user's message, but also a predefined system prompt with detailed rules for the AI assistant, and in some cases previous conversation context. Therefore, token usage does not directly reflect the length or complexity of the user's own input.

Overall, the collected dataset confirms that the implemented telemetry and AI metadata collection mechanisms are functional and provide sufficient information for descriptive analysis of user behavior and AI usage patterns.

3.2 AI Usage Patterns

The analysis of AI usage shows that users interacted with the AI assistant actively during the selected period. A total of 164 AI requests were recorded from 13 active users, resulting in an average of 12.6 requests per active user.

This indicates that AI usage was distributed across most users in the dataset rather than being generated by only one isolated account. However, request activity was not equal across all demographic groups. Some groups showed higher average AI request activity than others, suggesting differences in engagement intensity.

The age group 18–24 demonstrated the highest AI request activity, with 5 users generating 85 requests and an average of 17.0 requests per user. The age group 25–34 included 5 users and generated 52 requests, with an average of 10.4 requests per user. The 55+ group generated 15 requests among 2 users, while the 45–54 group generated 12 requests among 2 users.

These results suggest that younger users in the dataset interacted with the AI assistant more frequently than older users. However, due to the small group sizes, especially in the 45–54 and 55+ groups, this result should be interpreted as an observed pattern rather than a generalizable conclusion.

3.3 System Performance Analysis

System performance was evaluated using average response time and token usage. During the selected period, the average AI response time was 3424.0 milliseconds. This means that, on average, users received AI responses in approximately 3.4 seconds.

The average token usage was 3078.4 tokens per request. This value reflects the total amount of text processed by the AI model, including the user message, the predefined system prompt, assistant instructions, and possible conversation context. Therefore, token usage in this study should be interpreted mainly as a technical resource consumption metric rather than a direct measure of user prompt complexity.

Performance also varied between user segments. For example, the 25–34 age group had an average response time of 2924.8 milliseconds, while the 55+ group had a higher average

response time of 4632.5 milliseconds. Male users had an average response time of 3781.2 milliseconds, while female users had an average response time of 3000.1 milliseconds.

These differences may be influenced by variation in request complexity, token usage, or the small number of users in some groups. Therefore, they should be treated as descriptive observations rather than evidence of direct demographic influence on system performance.

3.4 Demographic-Based Analysis

Demographic segmentation was performed using age group, gender, weight bucket, and height bucket. The purpose of this analysis was to examine whether AI usage differed between user groups.

The gender-based analysis showed that male and female users were evenly represented in the dataset, with 7 users in each group. Male users generated 89 AI requests, with an average of 12.7 requests per user. Female users generated 75 AI requests, with an average of 10.7 requests per user. This means that the male segment showed slightly higher average request activity in the selected period.

Age-based analysis showed the highest average request activity in the 18–24 group. This group had 5 users, 85 total requests, and an average of 17.0 requests per user. The 25–34 group had 5 users, 52 requests, and an average of 10.4 requests per user. The 45–54 group had 2 users, 12 requests, and an average of 6.0 requests per user, while the 55+ group had 2 users, 15 requests, and an average of 7.5 requests per user.

Weight-based analysis showed the highest average request activity in the <60 kg bucket, with 22.0 requests per user. However, this group included only one user, so the result may reflect individual behavior rather than a broader segment pattern. The 75–90 kg group included 6 users and generated 82 requests, with an average of 13.7 requests per user, making it a more stable segment for interpretation.

Height-based analysis showed that users in the 175–190 cm bucket generated the highest total number of requests. This group included 9 users and produced 115 requests, with an average of 12.8 requests per user. The 160–175 cm group included 4 users and generated 48

requests, with an average of 12.0 requests per user. The <160 cm bucket included only one user and generated one request, which limits meaningful comparison.

Overall, the demographic analysis suggests that AI usage differed between user groups. The most visible differences were found in age and gender segments. However, because the dataset is small and some groups include only one or two users, the results should be interpreted carefully. The findings are useful for identifying preliminary patterns, but they do not establish statistically confirmed relationships.

4. DISCUSSION

4.1 Interpretation of Results

The results showed that AI usage varied across users and demographic groups. The most active age group was 18–24, while the 45–54 and 55+ groups showed lower average request activity. This may indicate that younger users in the dataset were more willing to interact with the AI assistant, although the small dataset size does not allow statistically validated conclusions.

The results also showed moderate gender-based differences. Male users generated slightly more AI requests on average than female users. However, because the total number of users was small, this difference should be interpreted as a preliminary observation rather than evidence of a strong demographic effect.

Token usage was relatively high, but this result should be interpreted in the context of the system architecture. Each AI request included a predefined system prompt and assistant instructions, meaning that token usage mainly reflects technical resource consumption rather than the length or complexity of the user’s own message.

Overall, the findings confirm that the implemented telemetry and analytics infrastructure can support descriptive analysis of AI usage, system performance, and demographic-based usage patterns.

4.2 Implications for System Design

Based on the observed usage patterns, the findings have several implications for the design of AI-powered fitness applications.

Personalization is likely to play an important role in improving user experience. Adapting AI responses based on user behavior and interaction history may help make the system more effective and relevant for different users.

Integration of AI features with core application functionality is also important. Embedding AI assistance directly into frequently used features, such as training or meal tracking, may increase adoption and usability.

In addition, improving the onboarding process may have a significant impact on long-term engagement. AI-based guidance during onboarding could help reduce friction and support users in understanding how to use the system effectively.

4.3 Limitations

This study has several limitations that should be considered.

First, the analysis is based on data collected from a single application, which may limit the generalizability of the findings.

Second, the analytical approach is primarily descriptive and does not include formal statistical methods such as regression analysis or hypothesis testing. As a result, the study focuses on identifying patterns rather than establishing statistically validated relationships.

Third, telemetry data collection follows a best-effort approach, especially on the client side. This means that some events may not be recorded due to network or technical constraints.

In addition, the central backend database stores only metadata related to AI interactions and does not persist full conversational transcripts. As a result, the thesis analysis is limited to quantitative aspects of usage and does not include deeper semantic analysis of user prompts or AI responses.

Finally, the available demographic variables are limited and may not capture all factors influencing user behavior.

4.4 Future Work

Although the collected dataset was sufficient for descriptive analysis, future work should include the collection and analysis of user data over a longer period. This would allow the validation of the analytical framework on a larger dataset and provide more stable empirical results.

Another important extension is the application of more advanced analytical methods. Techniques such as regression analysis, clustering, or AI model training could be used to identify deeper relationships between variables and to improve user segmentation.

Future work may also include the integration of semantic analysis of AI interactions. By analyzing the content of user queries and AI responses, it would be possible to better understand user intent and interaction quality.

Improving telemetry reliability is another important area. Implementing mechanisms such as event retry, offline buffering, or delivery guarantees would increase data completeness and improve the quality of analysis.

In addition, expanding the set of collected user attributes could provide a more comprehensive understanding of user behavior. This may include contextual or behavioral variables beyond basic demographic information.

Finally, future development may focus on adaptive AI systems that dynamically adjust their behavior based on user interaction patterns. Such systems could provide more personalized and effective support, improving both user experience and engagement.

CONCLUSION

This thesis focused on the analysis of user behavior and AI usage in a fitness application based on telemetry data. The main objective of the study was to explore how users interact with AI features and how these interactions may be related to demographic characteristics and engagement patterns.

To achieve this objective, a data-driven analytical framework was developed using the Gym-BRO application as a case study. The work included the design and implementation of a backend system, telemetry data collection mechanisms, a structured PostgreSQL data storage model, and analytical procedures for descriptive analysis, time-based aggregation, and demographic segmentation.

The collected dataset confirmed that the implemented analytics infrastructure is functional. During the selected analysis period from 18 March 2026 to 11 May 2026, the system recorded 14 users, 60 sessions, 910 telemetry events, and 164 AI requests. Out of 14 users, 13 interacted with the AI functionality, with an average of 12.6 requests per active user. The average token usage was 3078.4 tokens, and the average AI response time was 3424.0 milliseconds. Since each request included a predefined system prompt with detailed AI assistant instructions, token usage was interpreted as a technical resource consumption metric rather than a direct indicator of user prompt length or complexity.

The results showed that AI usage differed between user groups. The highest request activity was observed in the 18–24 age group, while the 45–54 and 55+ groups showed lower average activity. Gender-based analysis also showed a moderate difference, with male users generating slightly more AI requests on average than female users. Additional differences were observed across weight and height groups, although some of these segments contained only one user and therefore should be interpreted carefully.

The findings demonstrate that telemetry data can be used to identify preliminary usage patterns, evaluate AI system performance, and compare user activity across demographic segments. However, the analysis remains descriptive and limited by the small dataset size. Therefore, the results should not be interpreted as statistically validated relationships.

From a practical perspective, this work highlights the importance of integrating telemetry-based analytics into AI-driven applications. Such integration supports data-driven decision-making and creates opportunities for improving personalization, user experience, and system reliability.

Overall, the thesis contributes a system-oriented approach to studying AI usage in a real application environment. The developed framework can be extended in future work through

longer-term data collection, more advanced statistical analysis, improved telemetry reliability, and deeper feature-level analysis.

KOKKUVÕTE

Käesolev lõputöö keskendus kasutajakäitumise ja tehisintellekti kasutamise analüüsile treeningrakenduses telemeetriaandmete põhjal. Uuringu peamine eesmärk oli uurida, kuidas kasutajad suhtlevad tehisintellekti funktsioonidega ning kuidas need interaktsioonid võivad olla seotud demograafiliste tunnuste ja kasutusaktiivsuse mustritega.

Selle eesmärgi saavutamiseks töötati GymBRO rakenduse põhjal välja andmepõhine analüütiline raamistik. Töö hõlmas backend-süsteemi kavandamist ja rakendamist, telemeetriaandmete kogumise mehhanismide loomist, struktureeritud PostgreSQL andmesalvestusmudeli väljatöötamist ning analüütiliste protseduuride määratlemist kirjeldava analüüsi, ajapõhise agregeerimise ja demograafilise segmentimise jaoks.

Kogutud andmestik kinnitas, et rakendatud analüütikataristu on toimiv. Valitud analüüsiperioodil 18. märtsist 2026 kuni 11. maini 2026 registreeris süsteem 14 kasutajat, 60 sessiooni, 910 telemeetriasündmust ja 164 tehisintellekti päringut. Neljateistkümnest kasutajast 13 kasutasid tehisintellekti funktsionaalsust ning aktiivse kasutaja kohta tehti keskmiselt 12,6 päringut. Keskmine token'ite kasutus oli 3078,4 token'it ning keskmine tehisintellekti vastamisaeg oli 3424,0 millisekundit. Kuna iga päring sisaldas lisaks kasutaja sisendile ka eelnevalt määratud süsteemprompti koos detailsete juhistega tehisintellekti assistendile, tõlgendati token'ite kasutust selles töös eelkõige tehnilise ressursikasutuse näitajana, mitte kasutaja enda sisendi pikkuse või keerukuse otsese mõõdikuna.

Tulemused näitasid, et tehisintellekti kasutamine erines kasutajagruppide lõikes. Kõige suurem päringuaktiivsus ilmnas vanuserühmas 18–24, samas kui vanuserühmades 45–54 ja 55+ oli keskmine aktiivsus madalam. Soolise jaotuse põhjal ilmnas samuti mõõdukas erinevus: meessoost kasutajad tegid keskmiselt veidi rohkem tehisintellekti päringuid kui naissoost kasutajad. Täiendavaid erinevusi täheldati ka kaalu- ja pikkuserühmade lõikes, kuid osa neist segmentidest sisaldas ainult ühte kasutajat ning seetõttu tuleb neid tulemusi tõlgendada ettevaatlikult.

Tulemused näitavad, et telemeetriaandmeid saab kasutada esialgsete kasutusmustrite tuvastamiseks, tehisintellekti süsteemi jõudluse hindamiseks ning kasutajaaktiivsuse

võrdlemiseks demograafiliste segmentide lõikes. Samas jääb analüüs kirjeldavaks ning seda piirab väike andmestiku maht. Seetõttu ei tohiks tulemusi tõlgendada statistiliselt kinnitatud seostena.

Praktilisest vaatenurgast rõhutab töö telemeetrial põhineva analüütika integreerimise olulisust tehisintellektil põhinevatesse rakendustesse. Selline integratsioon toetab andmepõhist otsustamist ning loob võimalusi personaliseerimise, kasutajakogemuse ja süsteemi töökindluse parandamiseks.

Kokkuvõttes annab lõputöö süsteemipõhise lähenemise tehisintellekti kasutamise uurimiseks reaalses rakenduskeskkonnas. Väljatöötatud raamistikku saab tulevases töös edasi arendada pikemaajalise andmekogumise, keerukamate statistiliste analüüside, telemeetria töökindluse parandamise ja põhjalikuma funktsioonipõhise analüüsi kaudu.

REFERENCES

1-2. Mariani, M. M., Hashemi, N., & Wirtz, J. (2023). Artificial intelligence empowered conversational agents: A systematic literature review and research agenda. *Journal of Business Research*, 161, 113838.

Available at: <https://www.sciencedirect.com/science/article/pii/S0148296323001960>

3. Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science*, 11(6), 838-854

Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5572675/pdf/nihms862908.pdf>

4. Microsoft. (2026). Application Insights telemetry data model

Available at: <https://learn.microsoft.com/en-us/azure/azure-monitor/app/data-model-complete>

5. Qlik. What is descriptive analytics?

Available at: <https://www.qlik.com/us/reporting-analytics/descriptive-analytics>

6. Chew S, Abd Razak SS, Lo YT. The relationship between artificial intelligence usage and academic performance of students at a Malaysian public university. *Int J Pervasive Comput Commun*. 2025;ahead-of-print. doi:10.1108/IJPCC-05-2025-0186

Available at: <https://www.emerald.com/ijpcc/article/doi/10.1108/IJPCC-05-2025-0186/1317219/The-relationship-between-artificial-intelligence>

7. Manero A, et al. Profile segmentation: Clustering approach based on demographics and behavioral traits. *SAGE Journals*. 2025.

Available at: <https://journals.sagepub.com/doi/10.1177/18761364251343210>

8. Lawson-Body A. Anomaly identification through data visualization. *Issues in Information Systems*. 2023;24(4):161-175. Available at: https://iacis.org/iis/2023/4_iis_2023_161-175.pdf

9. OpenAI. (2026) OpenAI API documentation.

Available at: <https://developers.openai.com/api/reference/overview>

10. PostgreSQL Global Development Group. (2026). PostgreSQL documentation.

Available at: <https://www.postgresql.org/docs/>

11. React. (2026). React documentation.

Available at: <https://react.dev/reference/react>

12. MDN Web Docs. (2026). HTTP overview.

Available at: <https://developer.mozilla.org/en-US/docs/Web/HTTP/Reference>

13. MDN Web Docs. (2026). JSON.

Available at: https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/JSON

14. Dexie.js. (2026). Dexie documentation.

Available at: <https://dexie.org/docs>

15. Capacitor. (2026). Capacitor documentation.

Available at: <https://capacitorjs.com/docs>

16. GitHub. (2026). Git documentation.

Available at: <https://git-scm.com/docs>

17. GymBRO Analytics Dashboard [GitHub repository].

Available at: <https://github.com/ivan-puhachov/gymbro-analytics>

18. GymBRO Application [GitHub repository].

Available at: <https://github.com/Dennel04/gymbro>