# Implementation and Evaluation of a Movie Recommendation System

**Kangdi Wang** and **Erzhizhi Hu** and **Zhenzhi Liu** and **Ziyang Wang** and **Kaiyang Yan**
Heriot-Watt University, Edinburgh, United Kingdom
{kw2021, eh2020, zl2046, zw2030, ky2014}@hw.ac.uk

## Abstract

This project aims to develop a recommendation system implemented with a large language model (LLM) and associated methodologies, such as Low-Rank Adaptation (LoRA) fine-tuning and Retrieval-Augmented Generation (RAG). The system is designed to deliver personalized responses to user queries, demonstrating a significant advantage over traditional systems that rely solely on keyword indexing and produce rigid, robotic replies. By integrating these advanced techniques, the proposed system aims to enhance the accuracy, relevance, and adaptability of its recommendations, thereby offering a more tailored and effective user experience.

## 1 Introduction

In recent years, the rapid advancement of artificial intelligence, particularly in the domain of natural language processing, has opened new avenues for developing sophisticated recommendation systems. Traditional recommendation systems, which often depend on keyword indexing and predefined response templates, frequently fall short in delivering personalized and contextually relevant outputs. These limitations result in rigid, robotic interactions that fail to meet the diverse and dynamic needs of users.

To address this gap, this project proposes the development of a recommendation system powered by a large language model, enhanced with cutting-edge techniques such as Low-Rank Adaptation fine-tuning and Retrieval-Augmented Generation. By promoting the capabilities of LLMs and these complementary methodologies, the system aims to provide highly personalized and adaptive responses to user queries, marking a significant improvement over conventional approaches.

### 1.1 Aims

The primary aim of this project is to design and implement a recommendation system that utilizes a large language model as its foundation, augmented by LoRA fine-tuning and RAG techniques. Figure 1 shows the overall framework design of this project. The system aims to overcome the shortcomings of traditional recommendation frameworks by delivering responses that are not only accurate and relevant but also tailored to the individual preferences and contexts of users. Through this approach, the project intends to establish a new benchmark for personalization and flexibility in recommendation systems, contributing to the broader field of artificial intelligence and user-centered technology design.

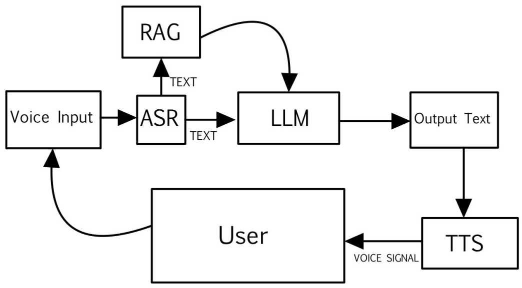



Figure 1: Project framework

### 1.2 Objectives

1. Integration of LLM and Advanced Techniques: To effectively incorporate a large language model with LoRA fine-tuning and RAG, ensuring seamless synergy between these components for optimal performance.

2. Personalization of Responses: To develop algorithms and mechanisms that enable the system to analyze user queries and generate responses customized to individual needs and preferences.

3. Evaluation of System Performance: To assess the system's effectiveness in terms of

accuracy, relevance, and user satisfaction, comparing its performance against traditional keyword-based recommendation systems.

4. Optimization and Scalability: To refine the system's architecture for efficiency and scalability, ensuring it can handle a diverse range of queries and a growing user base without compromising response quality.

## 2 Literature Review

The development of a dialogue-based movie recommendation system represents a significant advancement in the field of personalized entertainment. This literature review aims to provide a comprehensive overview of the existing research and methodologies related to the construction of such systems. By analyzing the relevant literature, we can identify key components, challenges, and opportunities in building an advanced movie recommendation system that takes advantage of voice interaction and large language models.

### 2.1 Recommendation Systems

#### 2.1.1 MovieLens

MovieLens is one of the most well-known and widely used recommendation system platforms in academia, developed by the GroupLens research team at the University of Minnesota. It provides multiple publicly available user rating datasets, serving as a standard benchmark for experiments and comparisons of movie recommendation algorithms. Early MovieLens recommendation algorithms were primarily based on collaborative filtering, particularly the analysis of user-item rating matrices and matrix factorization methods. With the advancement of research, deep learning models such as AutoRec have also been introduced to enhance recommendation performance. The openness and stability of MovieLens give it an important position in the research of movie recommendation systems (Harper and Konstan, 2015).

#### 2.1.2 Netflix

The Netflix recommendation system provides personalized viewing suggestions to hundreds of millions of users every day. The Netflix Prize competition held in 2006 propelled the development of matrix factorization, SVD, and hybrid recommendation algorithms, greatly advancing research in the field of recommendation systems (Gomez-Uribe and Hunt, 2016). Currently, the Netflix recommendation system employs a hybrid recommendation architecture that integrates content-based methods, collaborative filtering, context-aware recommendations, and extensively uses deep learning and graph models in engineering implementation. The system not only emphasizes the accuracy of the recommendations, but also values the balance between user engagement, recommendation diversity, and business metrics.

#### 2.1.3 Amazon Prime Video

The recommendation system of Amazon Prime Video introduces multimodal learning technologies, including text, images, and behavioral data, based on traditional collaborative filtering methods. For instance, Amazon employs a Transformer architecture similar to BERT in its research to perform sequence modeling, capturing users' long-term and short-term preferences from their historical viewing behavior, thereby enhancing the contextual relevance of recommendations. This deep sequence modeling approach is particularly suited for time-sensitive content consumption scenarios, such as movie and television recommendations (Sun et al., 2019)

### 2.2 Key Components

#### 2.2.1 Large Language Models (LLMs)

Large language models, such as GPT-3 and BERT, are pre-trained on vast amounts of text data and can generate human-like responses. In the context of movie recommendations, LLMs can be fine-tuned to understand user queries, provide relevant suggestions, and maintain context across multiple turns of conversation (Devlin et al., 2019; Radford et al., 2019). Fine-tuning is crucial to adapting the model to the specific domain of movies and user preferences.

#### 2.2.2 Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation is a hybrid approach that combines the strengths of retrieval-based and generative models. In a movie recommendation system, RAG can retrieve relevant movie information from a large database and generate coherent responses based on the retrieved context (Lewis et al., 2020). This approach ensures that the recommendations are accurate and contextually relevant. Data cleaning and enhancement play a vital role in improving the quality of the retrieved information.

#### 2.2.3 Voice Interface

Voice interfaces have gained popularity due to their ease of use and accessibility. Integrating voice recognition and text-to-speech (TTS) technologies allows users to interact with the recommendation system hands-free. Azure Automatic Speech Recognition (ASR) and other similar platforms provide robust solutions for voice interface integration (Karpukhin et al., 2020).

### 2.3 Challenges and Opportunities

#### 2.3.1 Context Management

In a conversation, users want the system to remember their preferences, previous queries, and the flow of the conversation. However, in dialogue based systems, maintaining contextual consistency and coherence across multiple rounds of dialogue is a challenge. Advanced techniques such as context embeddings and memory networks are being explored to address this challenge (Li et al., 2017).

#### 2.3.2 Hallucination

Unlike traditional information retrieval systems, LLMs generate responses based on probabilistic associations learned from vast corpora rather than structured knowledge bases. This often leads to outputs that include fabricated citations, incorrect definitions, or invented facts (Ji et al., 2023). To mitigate hallucination, several techniques have been proposed. Retrieval-Augmented Generation, which integrates external knowledge retrieval with text generation, has proven effective in improving factual consistency (Lewis et al., 2020).

## 3 Methodology

### 3.1 Data source

In this study, we selected the ReDial dataset to train our model, a resource widely recognized in the development of recommendation systems, particularly for movie content recommendations. The ReDial dataset, as documented by Li et al. (2018), comprises annotated dialogues wherein users engage in movie recommendation exchanges, offering a robust foundation for conversational recommendation research.

To prepare the data for our purposes, we first segmented the multi-turn dialogues into discrete dialogue pairs. Subsequently, we applied a filtering process to exclude samples containing invalid sentences, such as those that were blank, unrelated to movies, or characterized by an excessively strong tone—to ensure data quality and relevance. Following this preprocessing, we developed a script to reformat the filtered dialogue pairs into a structured JSON-like format: ‘ "instruction": "", "input": "", "output": ""‘, thereby aligning the data with the requirements of subsequent tasks in our methodology. This structured approach helps precise model training and evaluation within the context of our research objectives.

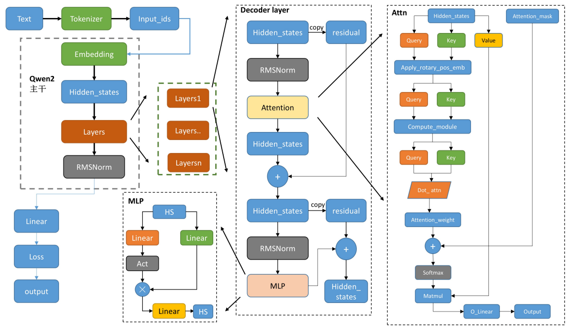



Figure 2: Qwen overview

### 3.2 LLM

We choose Qwen-2.5-7b-Instruct as our base model, it is a fintuned version of Qwen2.5-7b, with better command-follow abality(Team, 2024).

Qwen2 introduces several notable characteristics that contribute to its efficacy in this context. First, it applies Grouped Query Attention (GQA) across all model sizes, a mechanism that optimizes attention computation by grouping queries, thereby improving efficiency without compromising accuracy. This design choice ensures scalability and resource efficiency, making it suitable for deployment across varying computational constraints. Second, Qwen2 enables the bias option in the linear layers of the Query, Key, and Value (QKV) components within the attention mechanism. This adjustment allows for greater flexibility in modeling complex relationships within the data, potentially enhancing the model's representational capacity. Third, the model refines the operational logic of Rotary Position Embeddings (RoPE), optimizing how positional information is encoded and processed. This improvement leads to more effective handling of sequential data, particularly in tasks requiring precise contextual understanding over extended dialogues. Together, these features position Qwen2 as a robust and adaptable choice for our recommendation system.

### 3.3 RAG

In the development of our Retrieval-Augmented Generation system, we adopted the LangChain framework as the foundational architecture for integrating LLMs with external knowledge retrieval. This decision was informed by several distinct advantages offered by LangChain that align with the objectives of our system design. Firstly, LangChain provides a modular and extensible structure, enabling seamless integration of diverse data sources and retrieval mechanisms, which is critical for enhancing the contextual accuracy of the RAG system. Secondly, its built-in support for memory-augmented interactions allows the system to maintain coherence across multi-turn dialogues, a feature particularly beneficial for applications requiring sustained conversational context. These capabilities collectively contribute to improved performance in retrieving relevant information and generating informed responses. By applying LangChain, our system architecture ensures scalability, flexibility, and efficiency, positioning it well to address the challenges of real-time recommendation and knowledge-intensive tasks.

### 3.4 ASR and TTS

In our system, we employ OpenAI's APIs for both Automatic Speech Recognition and Text-to-Speech modules. The ASR component utilizes OpenAI's latest gpt-4o-transcribe model, which demonstrates enhanced accuracy and reliability, particularly in challenging scenarios involving diverse accents, background noise, and varying speech speeds. This model effectively transcribes audio into text with high precision.

For the purpose of speech synthesis, we incorporate OpenAI's GPT-4o-mini-TTS model, which provides enhanced fluency and dynamic tonal variation. This enables the generation of highly natural and human-like speech, with effective control over emotional features such as pauses, intonation, and speech rate. The synthesized speech exhibits a notable degree of emotional expressiveness, enhancing user immersion. Additionally, the TTS API provides low-latency generation, making it particularly suitable for real-time conversational applications.

## 4 Implementation

### 4.1 RAG

The acquisition and processing of movie data constitute a fundamental part of the RAG module implementation and represent the initial tasks in the project's implementation phase. Firstly, we obtained the original movie dataset through the free API provided by the OMBA. This dataset includes movie titles, genres, directors, actors, ratings, release years, and movie overviews. Then, we performed data cleaning on this dataset by writing dedicated scripts tailored to address various data issues, such as filling or deleting missing values, removing redundant data, and eliminating anomalies. For example, missing values in text fields (such as overviews, directors, and actors) were uniformly filled with empty strings, while numerical fields such as ratings were populated with averages or mode values to maintain data consistency. Additionally, regular expressions were used to standardize the data by removing special characters, non-English characters, and other anomalies to ensure uniformity in data formatting. During the data integration stage, we use another custom script to combine multiple movie-related fields (including title, genre, director, actors, and overview) into a unified comprehensive text field, facilitating subsequent text embedding operations. Simultaneously, numerical data such as ratings and release years were standardized to minimize adverse impacts from varying data scales on subsequent embedding results.

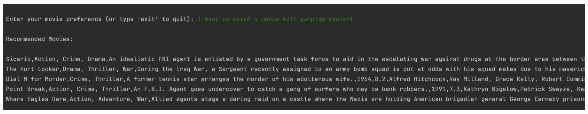



Figure 3: Output of relevance search

After data cleaning and integration were completed, we employed the Sentence Transformers model to generate text embeddings for the integrated movie information (including title, genre, director, actors, and overview). This process converted textual information into high-dimensional semantic vectors, effectively capturing semantic similarities among movies. Next, we used FAISS to build an efficient index of these semantic vectors, significantly enhancing retrieval efficiency and enabling the system to quickly and accurately find movies that closely match user queries at the semantic level. Additionally, we implemented auto-

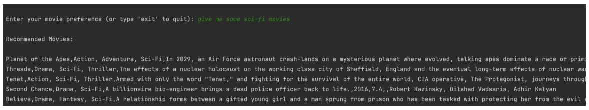



Figure 4: Output of precise search

matic parsing of user inputs using the spaCy natural language processing tool, specifically extracting precise search attributes such as genre, director, actor, and release year directly from the user's natural language input.

In our recommendation system, we combine two search strategies, namely precise search and relevance search. First, precise search is based on structured fields within movie data, such as genre, director, actor, and release year. Using a natural language processing tool (spaCy), these attributes are automatically extracted from the user input, precisely filtering a subset of movies that accurately match explicit user requirements. This method ensures the rapid retrieval and return of relevant movies from the RAG when the input requires quick lookup based on specific attributes. Then, we further apply a relevance search (semantic retrieval) to the subset of movies obtained from the precise filtering. We use the Sentence Transformers model to transform the combined textual fields of movies (including title, genre, director, actors, and overview) into semantic embedding vectors, and construct an efficient semantic index with FAISS. System can quickly and accurately identify movies from the precisely filtered subset that have the highest semantic similarity to the user's input, returning results sorted by relevance from highest to lowest.

The form of output strictly adhering to the format of the original movie database CSV file, including movie title, genre, overview, release date, rating, director, and actors. Each recommended movie's information is separated by commas to maintain data consistency and readability. This standardized output format facilitates direct user viewing and enables the subsequent language model to use the structured output from the RAG as context inputs, further generating personalized recommendation explanations, detailed movie analyses, or other user-friendly recommendation content.

### 4.2 LLM

To ensure efficient operation in a local environment, we selected the Qwen-2.5-7B-instruct model, which offers a balanced parameter scale, as our base model. To enable the model to function as a


Figure 5: The original dataset (up) and the cleaned dataset (down)

movie recommendation assistant capable of structured interactions with users, we fine-tuned it using the LoRA method.

Our training and testing datasets were primarily sourced from several .jsonl files related to movie recommendation dialogues found on GitHub. After thorough cleaning and restructuring, the overall data quality and utility were significantly improved.

Given the large dataset and the high parameter count of the model, full training would have required substantial computational resources. To mitigate this, we employed 4-bit quantization to reduce memory usage and accelerate training. As such, we conducted only approximately 0.166 epochs of training. Despite the limited training cycles, the model's loss decreased from an initial 5.0212 to 0.0777, indicating that the model was effectively learning from the data.

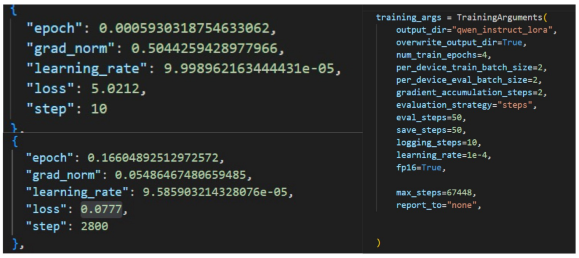



Figure 6: Fine-tuning results (left) and fine-tuning hyperparameters (right)

To guide the model's behavior more precisely during inference, we designed a dedicated system prompt (see Appendix A) that explicitly defines the model's role as a movie recommendation assistant. The prompt includes instructions on how to respond to user queries, and how to structure its output. This prompt is injected at the beginning of each dialogue session and plays a critical role in aligning the model's responses with the intended functionality, especially in scenarios where retrieval-based answers are not available. The use of a well-crafted prompt significantly enhances the consistency, tone, and relevance of the generated replies.

Furthermore, by incorporating RAG within a custom chat_loop function, we enabled the model to retrieve and leverage information from a local database in real time, allowing it to engage in personalized and context-aware conversations based on the retrieved content.


Figure 7: Answers are generated based on local database retrieval

Lastly, we deployed the model through the Ollama platform, which enabled efficient local execution and streamlined integration with other components in our system. By using OpenAI's API interfaces for Automatic Speech Recognition and Text-to-Speech technologies, we further extended the model's capabilities to support real-time voice-based interactions with users. This multimodal setup allows users to communicate with the recommendation assistant not only via text but also through natural speech, enhancing accessibility and user engagement in more diverse application scenarios.

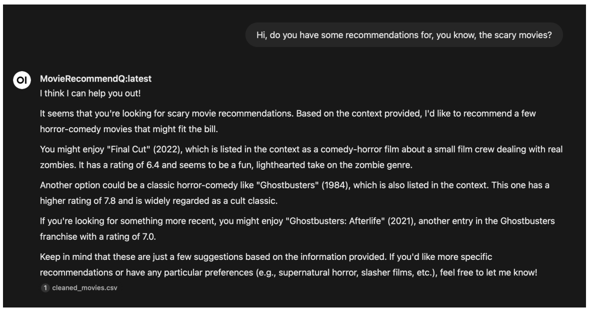



Figure 8: Deploy on Ollama

However, due to the limited size of the model, it still exhibits certain shortcomings. For instance, when users inquire about movies that are not present in the local database, the model does not explicitly indicate that no relevant information was found. Instead, it tends to generate plausible but hallucinated responses based on its pretraining knowledge, which may lead to misinformation or user confusion in specific contexts.

## 5 Evaluation and Results

### 5.1 Evaluation Experiment Design

To assess the performance and usability of our movie recommendation system, we designed a two-part evaluation framework. The first part focuses on measuring the semantic retrieval capability of the RAG module in isolation and the second part evaluates the overall system from a user-centric perspective through real interactions and subjective feedback. This dual approach allows us to separately analyze core functional modules and end-user experience.

#### 5.1.1 Evaluation for RAG

We first conducted an evaluation of the RAG system. Since precise search does not require accuracy evaluation, we focused primarily on the evaluation of approximate semantic matching. Specifically, we manually constructed a set of natural language queries that describe movies using combinations of genre, director, actor, and release year (e.g., "an action movie directed by Christopher Nolan"). For each query, we specified a target movie as the ground truth based on human judgment. The system then used FAISS and Sentence Transformers to return the top five most semantically relevant results. To assess retrieval performance, we examined three aspects: whether the target movie was ranked first (Top-1 accuracy), whether it appeared within the top three results (Recall@3), and whether it appeared anywhere in the top five (Recall@5). This setup allowed us to evaluate the quality of semantic matching independently of the language model generation component.

#### 5.1.2 Evaluation for System

To evaluate the system, we conducted a user study in which volunteers interacted with the system and completed a custom-designed questionnaire inspired by the System Usability Scale (SUS). The questionnaire included six items, each rated on a 5-point Likert scale, with scores ranging from 1 (far left) to 5 (far right). The first five items measured user satisfaction across the following dimensions: whether the system effectively addresses users' primary needs, comfort and efficiency in interacting with the system, quality and relevance of the recommended content, smoothness and naturalness of language during interactions, and sufficiency and usefulness of the explanatory details provided. The sixth item prompted users to compare our system with other commercial large language models and indicate their preference. To facilitate a fair and meaningful comparison, we also provided volunteers with access to the APIs of more advanced models, such as ChatGPT-4, allowing them to experience these models' movie recommendation capabilities. For evaluation, we calculated the average score for each question and each participant to assess overall usability.

### 5.2 Results and Analysis

In this section, we present the results of both our technical and user-based evaluations. We first analyze the effectiveness of the RAG module based on quantitative metrics, followed by an analysis of user feedback collected through questionnaires. This comprehensive analysis highlights both the strengths and limitations of the current system implementation, and offers insights into possible future works.

#### 5.2.1 Results and Analysis of RAG

The evaluation was conducted on 50 query-answer pairs. As shown in the bar chart below, in 28 out of the 50 cases, the target movie was ranked first, resulting in a Top-1 accuracy of 0.56. In 44 cases, the target movie appeared within the top three returned results, corresponding to a Top-3 accuracy of 0.88. In 46 cases, the target movie was successfully retrieved within the top five results, achieving a Recall@5 score of 0.92. These results indicate that the semantic embedding and vector indexing mechanisms are generally capable of accurately understanding user input and effectively ranking relevant results. The returned results reliably include the correct answer, and in more than half of the cases, the target movie is presented as the top result. RAG also plays an extremely important role in this project—not only providing accurate retrieval outputs for the subsequent LLM module, but also laying a solid foundation for the overall functionality and performance of the system.

#### 5.2.2 Results and Analysis of System

We recruited a total of 10 volunteers for this study. After analyzing their feedback, we found that the system's average usability score was 3.14 out of 5. Furthermore, aside from Question 3 (quality and relevance of the recommended content) and Question 5 (sufficiency and usefulness of the explanatory details), both of which received average scores below 3, all other items scored above 3. These

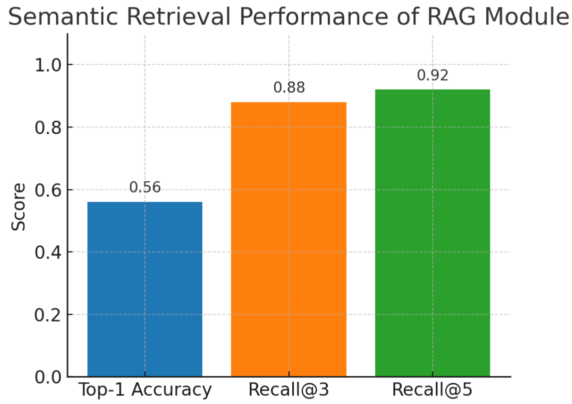



Figure 9: Semantic Retrieval Performance

findings suggest that while the system performs reasonably well in terms of addressing user needs, ease of interaction, and natural communication, there is still room for improvement in the content quality and the informativeness of the recommendations. In addition, for Question 6, which asked users to compare our system with ChatGPT-4, three participants expressed a preference for our system.

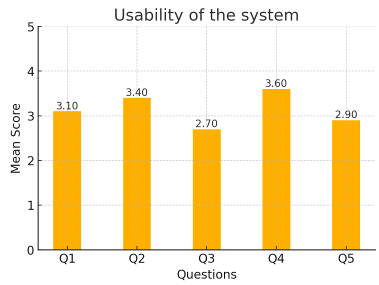



Figure 10: Usability scores of the system

The average scores for Questions 3 and 5 were below 3, indicating that most participants were dissatisfied with the variety of movie types, styles, or themes recommended by the system, as well as the depth of explanation provided for the recommended movies. We believe there are two main factors contributing to these lower scores. First, our RAG approach occasionally fails to precisely retrieve relevant content from the local database based on user input. Second, because the model itself has a relatively small parameter size and was only fine-tuned via LoRA, it sometimes struggles to correctly process and integrate the information returned by RAG—especially when the prompt includes a large amount of data, from both user and system sources. Consequently, the model may ignore or misinterpret certain instructions and recommend content that is not present in the local database. However, given that Question 1, which also depends on the RAG technique, received a relatively higher score, we are inclined to attribute these issues more to the model's limitations in synthesizing retrieved information rather than to the retrieval process itself.

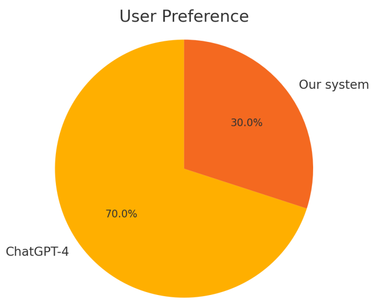



Figure 11: User preference

For Question 1, the average score was above 3, indicating that most participants felt the system's recommendations aligned with their expectations. This supports the conclusion that our system is capable of leveraging RAG techniques to retrieve relevant content from the local database based on users' descriptive inputs. For Question 2, the feedback suggests that most participants found the system relatively convenient and easy to use. This reflects the effectiveness of our overall system design, which integrates the Ollama platform with ASR and TTS technologies, enabling smooth voice-based interaction and seamless user experience. The results for Question 4 demonstrate that our fine-tuned language model, in conjunction with RAG output, can generate relatively natural and fluent responses. While the content may have certain limitations in depth or accuracy, the model is nonetheless able to deliver language that feels coherent and human-like in interaction. Finally, the fact that three participants indicated a preference for our system over ChatGPT-4 in Question 6, suggests that a locally deployed, lightweight model, when equipped with retrieval mechanisms and optimized for domain-specific tasks, can offer a competitive alternative to large-scale commercial language models, particularly in scenarios requiring offline access, data privacy, or customization.

## 6 Conclusion

In this study, we successfully developed a voice-interactive, locally deployed movie recommendation system by integrating Retrieval-Augmented Generation with a fine-tuned Qwen-2.5-7B-Instruct model hosted on the Ollama platform, alongside APIs for ASR and TTS technologies. We subsequently evaluated our system through performance experiments of the RAG and a questionnaire-based user study. Compared to systems relying on cloud-based APIs of larger language models, our system places greater emphasis on localization, lightweight deployment, data privacy, and flexibility for personalized or domain-specific customizations, making it particularly suitable for offline or resource-constrained scenarios.

Our study has several notable limitations. First, due to the constraints of local deployment, we were limited to selecting language models with relatively small parameter sizes. As a result, the recommendations provided by our system may sometimes lack depth and breadth, fail to accurately interpret system prompts, or even produce hallucinations. Second, hardware constraints restricted our fine-tuning process to Low-Rank Adaptation, rather than the more effective supervised fine-tuning approach, and also limited the number of training epochs. Consequently, the model occasionally struggles to correctly process and integrate the information retrieved by RAG, impacting overall performance and recommendation quality. Additionally, our evaluation methodology itself was relatively limited, as we relied solely on a questionnaire-based survey to evaluate system usability. Other important evaluation methods, such as in-depth interaction log analyses or objective performance benchmarks, were not employed in this study. This limits our ability to fully capture the system's strengths and weaknesses from multiple perspectives.

Future work could focus on exploring and comparing the performance of systems based on different locally deployable language models or evaluating how different fine-tuning strategies impact system effectiveness. Additionally, further research may consider employing model pruning techniques to reduce computational resource requirements, enabling the deployment of more sophisticated large language models on local hardware without compromising significant performance.

## Ethics Statement

AI ethical principles are moral norms guiding the design, development, deployment, and use of artificial intelligence. For instance, the "Ethical Guidelines for Trustworthy AI" (Commission et al., 2019) proposed by the European Commission stipulates that AI systems must ensure transparency, interpretability, and fairness during their design and deployment processes. In practice, AI ethics requires developers and users to prioritize data privacy and fairness, avoiding discriminatory outcomes caused by algorithmic bias. For example, training models with diversified datasets and conducting regular reviews and adjustments to the models can effectively prevent discrimination based on factors such as race and gender. Furthermore, AI ethics emphasizes the importance of attribution of responsibility and accountability mechanisms. Throughout the lifecycle of AI systems, from design to deployment, responsibilities must be clearly defined at every stage to ensure traceability of responsibility in case of issues.

Privacy protection is essential to ensure compliance with regulatory standards and maintain the integrity of scientific research. In the context of frameworks such as the EU General Data Protection Regulation (GDPR) and the UK Data Protection Act, researchers are obliged to strictly anonymize datasets by removing direct identifiers and generalizing indirect details. To reduce the risk of data misuse or discriminatory analysis, third-party platforms must adopt strong encryption technology. Post-study protocols require permanent deletion of raw data (including cloud backups) to prevent residual vulnerabilities. It is crucial to recognize that a decrease in trust due to privacy issues may directly undermine the credibility of research by reducing participation rates and compromising data validity.

Ethical research must be based on voluntary participation and ensure the avoidance of coercion through clear consent procedures. Uncontrolled bias in training data may marginalize minority groups. To address this issue, fairness indicators and debiasing techniques should improve models. Transparent design can cultivate user trust. Prioritizing fairness is not only a legal obligation but also a fundamental element of inclusive technology and a social obligation that every researcher should fulfill.

## References


European Commission, Content Directorate-General for Communications Networks, Technology, and Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji. 2019. *Ethics guidelines for trustworthy AI*. Publications Office.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Carlos A. Gomez-Uribe and Neil Hunt. 2016. The netflix recommender system: Algorithms, business value, and innovation. *ACM Trans. Manage. Inf. Syst.*, 6(4).

F. Maxwell Harper and Joseph A. Konstan. 2015. The movielens datasets: History and context. *ACM Trans. Interact. Intell. Syst.*, 5(4).

Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of hallucination in natural language generation. *ACM Comput. Surv.*, 55(12).

Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. 2020. Dense passage retrieval for open-domain question answering. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6769–6781, Online. Association for Computational Linguistics.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33:9459–9474.

Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky. 2017. Adversarial learning for neural dialogue generation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2157–2169, Copenhagen, Denmark. Association for Computational Linguistics.

Raymond Li, Samira Ebrahimi Kahou, Hannes Schulz, Vincent Michalski, Laurent Charlin, and Chris Pal. 2018. Towards deep conversational recommendations. In *Advances in Neural Information Processing Systems 31 (NIPS 2018)*.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. *OpenAI*. Accessed: 2024-11-15.

Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of* *the 28th ACM International Conference on Information and Knowledge Management*, CIKM '19, page 1441–1450, New York, NY, USA. Association for Computing Machinery.

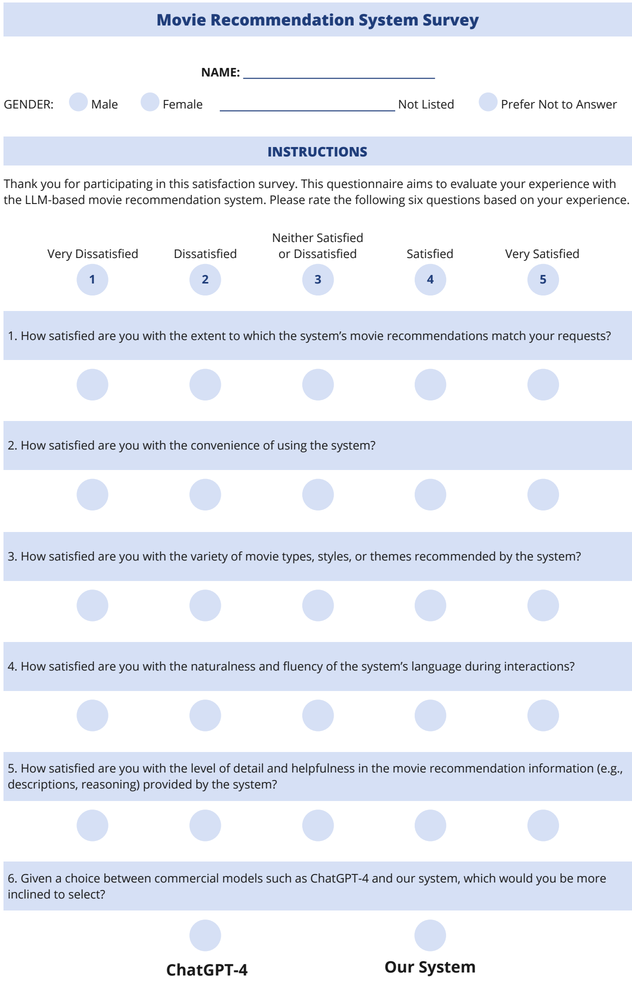
Qwen Team. 2024. Qwen2.5: A party of foundation models.

## A System Prompts

You are an intelligent movie recommendation assistant. Provide precise recommendations by combining user needs with knowledge base retrieval results. Follow this structured process:

1. **Output Format** Use this XML template in <recommendations>:
<title>"Movie Title (Year)"</title>
<rationale>Based on your interest in [keyword], this film [specific match]. Notably [unique feature]</rationale>
<attributes>
- Genre: Action/Sci-Fi
- Director: Christopher Nolan
- Keywords: Time travel, Mind-bending
</attributes>
</recommendation>
2. **Demand Analysis**
Classify user input into one of three categories:
- Type A: Specific parameters ("Recommend Hitchcock-style thrillers")
- Type B: Vague preferences ("Movies for couples")
- Type C: Direct references ("Similar to Inception")
3. **Information Matching**
Conduct in <analysis> section:
a) Extract key elements from user request
b) Cross-reference with movie database fields:
- Director/actor style
- Genre tags
- Keyword descriptors
- Douban similar recommendations
- Award information
4. **Recommendation Generation**
Adhere to these rules:
- Maximum 5 recommendations, ordered by relevance
- Each recommendation must include:
* Matching criteria ("Aligns with your sci-fi preference")
* Unique selling point ("Won Oscar for Best Visual Effects")
* Content advisory ("Contains mild violence")
- If no matches found, suggest refining criteria
- For vague requests, make knowledge-based inferences
Guidelines:
1. Only recommend movies present in knowledge base
2. Clearly distinguish objective data (ratings) from subjective terms ("mind-bending")
3. Include advisories for potentially sensitive content
4. Address conflicting requests ("light-hearted horror") in response preamble
5. Do not ask the user any questions like: "Do you want me to list some titles first or do you have any specific ones that you like already?"
Begin processing request.

# B Questionnaire and Feedbacks



**Movie Recommendation System Survey**

NAME: ____________________

GENDER: Male Female ____________________ Not Listed Prefer Not to Answer

**INSTRUCTIONS**

Thank you for participating in this satisfaction survey. This questionnaire aims to evaluate your experience with the LLM-based movie recommendation system. Please rate the following six questions based on your experience.

| Very Dissatisfied | Dissatisfied | Neither Satisfied or Dissatisfied | Satisfied | Very Satisfied |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

1. How satisfied are you with the extent to which the system's movie recommendations match your requests?

2. How satisfied are you with the convenience of using the system?

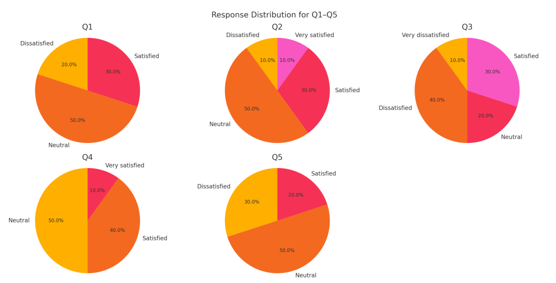
3. How satisfied are you with the variety of movie types, styles, or themes recommended by the system?

4. How satisfied are you with the naturalness and fluency of the system's language during interactions?

5. How satisfied are you with the level of detail and helpfulness in the movie recommendation information (e.g., descriptions, reasoning) provided by the system?

6. Given a choice between commercial models such as ChatGPT-4 and our system, which would you be more inclined to select?

ChatGPT-4 Our System



## C Consent Form

**Consent Form:** Take a Part in the CA model Evaluation
**Principal Investigator:** Hu Erzhizhi
**Contact:** eh2020@hw.ac.uk
**Objective of the experiment:**
This research aims to investigate the usability of a movie recommendation system and conduct a simple comparison with other commercial models.

**Procedure:**

- Try our movie recommendation system and another one based on ChatGPT-4.
- Complete the questionnaire.
- Allocate approximately 15 minutes for the entire session.

**Voluntary Participation:**
Participation is entirely voluntary. You have the right to withdraw at any point without penalty or consequences.

**Confidentiality and Data Use:**
All collected data will be anonymized to ensure the protection of your privacy. The data will be used solely for academic purposes, such as publications and presentations, with no personally identifiable information disclosed.

**Potential Benefits and Risks:**
The benefit is that you can use our model, and also access our ChatGPT-4 API for free during the testing period.

This test is risk-free.

**Consent:**
By signing below, you confirm that:

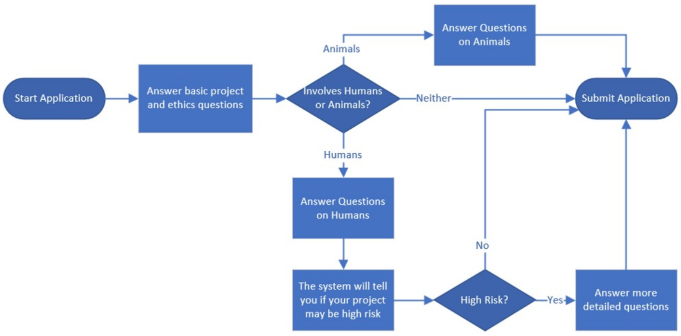
(1) You have read and understood this consent form.
(2) You voluntarily agree to participate in this experiment.
(3) You understand your right to withdraw at any time without consequences.
(4) You consent to the anonymized use of your data for academic purposes.

| | |
|---|---|
| **Name:** | **Signature:** |
| **Date:** | **E-mail address:** |

# D Ethic Form

## Ethics Application Form

### What is Going to Happen?




To complete an Ethics Application you will first answer some basic questions, followed by more specific questions depending on whether you are using human participants, personal data from external sources, or living animals.

"Part 1 - Ethics" asks you about ethical considerations your project may raise, and what you will do to address these ethical issues.

The section "Part 2 - Data Protection" will be asked if you are using human participants, or personal data from external sources. It will ask you about how you are using individuals data, and how you are protecting it. If your project is identified as potentially high risk you will need to answer further questions.

### Eligibility to Request Ethics Approval

The Heriot-Watt Research Ethics Policy can be found here. Before proceeding with your application please check to confirm the following:

☑ I have read and understood the Heriot-Watt Research Ethics Policy

### How to Navigate the System using the Left Panel:

| | | | |
|---|---|---|---|
| To move to the next page of the application: Next | To skip to different sections: Navigate | To save the application at any time: Save | To submit your application: Submit |

## Project Information

Type of research project: Undergraduate/Taught Postgraduate (MSc, MBA etc.)

Research based at: Edinburgh

School research is for: Mathematical and Computer Sciences

Department: Computer Science

## Applicant Details

Student Details

First Name: Kaiyang

Surname: Yan

Email: ky2014@hw.ac.uk

## Supervisor

Supervisor Details

First Name: Matthew

Surname: Aylett

School: MACS

Campus: Edinburgh

Telephone:

Email: m.aylett@hw.ac.uk

## Project Summary

Project title:

Movie Recommendation System Survey

## Ethical Considerations Check

Does your project involve any of the following? (Please check all that apply)

- [x] Human participants
- [ ] Personal data from external sources
- [ ] Living animals
- [ ] Medicines/Drugs/Medical Appliance
- [ ] None of the above

## Type of Approval

Type of approval requested: Full Approval

**Consider the following type of research project:**

Research projects that only involve use of data-sets obtained from external sources where the data is publicly available and was already truly anonymised when obtained from that source.

The data-set obtained is representative and inclusive of the target group and will not reinforce any bias or inequality.

Does the research project you are requesting ethical approval for **only** involve undertaking an activity defined above?

- ( ) Yes
- (•) No

## External Approval

Is ethical approval required by another body linked to the research, e.g. the NHS or a collaborator?

- ○ Yes
- ◉ No

## Details

State the question to be answered and the value of answering it (include the aims of the project):

In this project, we aim to investigate whether a data-driven film recommendation service can effectively meet the different needs of users, and why some people would choose not to use such a system. Analysing user interactions and feedback through large-scale data models, we seek to identify the strengths and limitations of current recommendation algorithms in meeting user preferences, as well as the barriers that prevent non-users from adopting the service (e.g., lack of trust, perceived irrelevance, or privacy concerns).

The value of answering these questions lies in leveraging insights to refine recommendation models, improve personalisation accuracy and increase user satisfaction. By addressing gaps in service delivery, the project aims to provide actionable recommendations for developers and platforms to increase adoption, promote inclusivity (e.g., accommodating underrepresented populations) and ensure ethical data practices.

Ultimately, this research aims to improve the fairness and effectiveness of AI-powered entertainment services while prioritising user privacy and transparency.

## Method(s)

Please state what methods or procedures will be used to collect and analyse data.

Use specific names for relevant details: If you are surveying, specify how you will be surveying *(e.g. on SurveyLab, Microsoft Forms, JISC Online Surveys)*, or if you are collecting data from Social Media specify which Social Media *(e.g. Facebook, Twitter)*

The team creates a questionnaire and sends it to the testers using WeChat. Feedback on the questionnaire is collected for the survey.

## Previous Experience

Does the researcher(s) involved in the project have relevant training, or previous experience *(e.g. have previously researched or collected data)* in the field of investigation?

- ○ Yes
- ◉ No

## Conflict of Interest

A "conflict of interest" is any case where the researcher(s) could profit in any way; personal or department, financial or otherwise, or if there is a pre-existing dependent relationship between a researcher and a participant (such as manager/employee, staff/student, client/consumer)

It is important to be transparent about any potential conflicts of interest to ensure any existing relationships are not conflicted by participation/non-participation.

Is there any potential "conflict of interest" relating to the proposed research project?

- ○ Yes
- ◉ No

## Confidential Information

Does the project involve using confidential information that is not already in the public domain? *(This does not include the confidential information of e.g. a research participant, but considers external information for example about a company which should be used appropriately without confidential information being divulged)*

- ○ Yes
- ◉ No

## Duration

State the likely duration of the project:

(Please note that the data collection can only commence following ethics approval)

One semester.
Data collection and analysis will take place in weeks 7-12 of Semester 2,2025

## Location

In which country, or countries will the research take place? *If research is online select "Worldwide"*

United Kingdom

On which premise(s) or location(s) will data be collected?

If the project is to be undertaken online state this here and provide details. *E.g. if you are monitoring social media, specify which social media you are monitoring.*

Distribute the questionnaire online using WeChat. Collecting information through WeChat.

## Human Participants

State the type of participant(s) who will be involved?

Univercity MSc students on HWU course F21CA

How many individuals will participate in the research? (State the maximum foreseen number).

10

## Vulnerable Participants

Will any participants be from any of the following vulnerable groups?

- [ ] Children
- [ ] People with learning disabilities
- [ ] Patients in hospital
- [ ] Participants with mental health issues
- [ ] Other (e.g. homeless people, refugees, people who lack capacity to consent etc.)
- [x] N/A - Participants are not from any vulnerable groups

## Person Collecting Data

Who will collect the information from the participants?

The research team will distribute and collect survey data via a microsoft questionnaire. Designated team members will administer participant responses, ensure anonymity, and store data securely. No third parties will be allowed to access the raw data without consent.

### Non-Standard Hardware

Will participants be using specialist hardware? *For example eye-trackers or development prototypes?*

- ○ Yes
- ◉ No

### Physical Hazards

Are there any other potential physical hazards to participants including personal security?

- ○ Yes
- ◉ No

### Recruitment

State how and where participants will be recruited. Be specific, if you are recruiting via Social Media please reference each Social Media Platform you will use:

WeChat: Expanded coverage and sought out film enthusiasts to participate in the test through the research team's use of circle-of-friends retweets.

How long will a participant have to decide whether to take part in the research?

Standard research time:
The complete process (questionnaire) usually takes 1 day per participant to ensure adequate communication and data integrity.

Opportunistic contact handling:
In the event that a potential participant is encountered in an unplanned scenario (e.g., between other interviews), the research team will briefly explain the purpose of the study, voluntariness, and privacy terms (an electronic briefing page will be provided).
Participants can decide on the spot whether to join or not, which is done via a WeChat code.
Even if the participant chooses to enrol immediately, the team will follow up with a full study description document to ensure that he/she is fully informed.

Ethical safeguards:
No pressure or inducement to participate is permitted under any circumstances, and withdrawal is clearly understood to be possible at any time and without consequence.
Emergency contact and complaint channel information is provided with the consent form.

Can you provide any copies of advertisements/recruiting matter you will be providing to participants?

- ○ Yes
- ◉ No

## Compensation for Participation

Will compensation be provided to participants? (Financial or otherwise)

- ○ Yes
- ◉ No

## Consent to Participate

Will informed consent to participate be obtained from all appropriate parties?

- ◉ Yes
- ○ No

How will you be obtaining informed consent to participate?

- ☐ Written consent
- ☑ Audio / verbal consent
- ☑ Electronic consent, e.g. via online survey
- ☐ Other

Upload the wording (or if exact wording cannot be supplied, suggested wording) that will be used for verbal recorded consent here:

Documents

| Type | Document Name | File Name | Version Date | Version | Size |
|---|---|---|---|---|---|
| Consent Form | wording | wording.pdf | | | 42.2 KB |

Save the electronic consent page as a .pdf and upload here:

Documents

| Type | Document Name | File Name | Version Date | Version | Size |
|---|---|---|---|---|---|
| Consent Form | Consent Form | Consent Form.pdf | | | 133.2 KB |

## Informing Participants

How will you tell individuals about the project, the use of their data, and who to contact if they want to find out more? (Select all that apply)

- [x] Privacy Notice
- [x] Participant Information Sheet
- [x] Plain Language Statement
- [ ] Debrief Form
- [ ] Other
- [ ] N/A - Individuals will not be told about the project

Upload your Privacy Notice here:

Documents

| Type | Document Name | File Name | Version Date | Version | Size |
|---|---|---|---|---|---|
| Privacy Notice | Privacy Notice | Privacy Notice.pdf | | | 45.9 KB |

Upload your Participant Information Sheet here:

Upload your Plain Language Statement here:

Documents

| Type | Document Name | File Name | Version Date | Version | Size |
|---|---|---|---|---|---|
| Privacy Notice | Plain Language Statement | Plain Language Statement.pdf | | | 61.1 KB |

## De-Identification Procedures

**Pseudonymisation** is a de-identification procedure where personally identifiable information is replaced by an alternative identifier, or “pseudonym”. When data is pseudonymised, a person can be re-identified by that data, but only with the use of additional data, or a ‘key’. (Sometimes pseudonymised is known as "linked anonymised")

**Anonymisation** is a de-identification procedure whereby once personal data is completely anonymised it can never be used to re-identify a person (even in combination with other data).

Will you be using any pseudonymisation techniques or procedures?

- ( ) Yes
- (•) No

Will you be using any anonymisation techniques or procedures?

- ( ) Yes
- (•) No

### Emotional Discomfort or Distress

Will the project involve procedures that may cause emotional discomfort or distress to participants which may have long lasting or significant effects?

- ○ Yes
- ◉ No

### Deception

Will the project involve deceiving a participant or providing incomplete disclosure?

- ○ Yes
- ◉ No

Will the project involve a deception or incomplete disclosure which could have any long lasting or significant effects on the participant?

- ○ Yes
- ◉ No
- ○ N/A

### Mandatory or Voluntary?

Is participation in this project voluntary? *Participation may not be voluntary, for example if a participant does not know that they are being observed.*

- ◉ Yes - Participation is voluntary
- ○ No - Participation is not voluntary

### Privacy Intrusive

Might the project require you to contact individuals in ways they may find intrusive to their privacy, and that may have a long lasting or significant impact on them?

- ○ Yes
- ◉ No

### General Practitioner, Medical Specialist or Family Doctor Informed?

If the research may have an adverse impact on the physical or mental health of a participant, will the participants' Medical Specialist, General Practitioner or Family Doctor be informed of the recruitment of the participant before the research project begins?

(This includes any medical practitioner of whom the participant is a patient).

- ○ Yes
- ○ No
- ◉ N/A - There is no need to inform

## Types of Data Collected and Analysed

Specify the categories of personal data to be collected and analysed: (Select all that apply)

- ☐ Political Opinions
- ☐ Religious or Philosophical Opinions
- ☐ Racial or Ethnic Origin
- ☐ Trade Union Membership
- ☐ Physical Health
- ☐ Mental Health
- ☐ Sex Life
- ☐ Sexual Orientation
- ☐ Alleged Offences or Proven Offences
- ☐ Gender Identity
- ☑ Other (e.g. name, age range, location, interactions, opinions etc.)

If "Other" specify the other categories of personal data:

## Potential Danger to Individuals

Will you be collecting or processing data that might endanger the individual's physical health or safety in the event of a security breach?

- ○ Yes
- ◉ No

## Part 2 - Data Protection

You have now completed **Part 1 - Ethics**. The next set of questions will consider how the project plan incorporates data protection by design.

## Data Protection and Privacy Notices

You were previously asked about how you would tell individuals about the project, the use of their personal data, and who to contact if they want to find out more *(for example a Privacy Notice, Participant Information Sheet, Plain Language Statement or Debrief Form).*

---

Can you confirm that the given method of telling individuals about the nature and purpose of the project, either;

- Uses a template or wording approved by the Data Protection Officer; or

Includes all of the following:
- Data Protection Officer contact details: dataprotection@hw.ac.uk
- A link to the Heriot-Watt University Privacy Notice for Research Participants (or how to find it)
- Information on which party is the data controller

(•) Confirm
( ) Cannot Confirm

## Privacy Invasive Technology or Algorithms

Will your research involve use of any technology or algorithm which may be perceived as being privacy intrusive, and may have a long lasting or significant effect on individuals? *For example consider any algorithms which may embed bias or discrimination?*

( ) Yes
(•) No

## Systematic Monitoring

Will your research involve any systematic monitoring, which would include processes which observe, monitor or control individuals, and may have a long lasting or significant impact on individuals?

( ) Yes
(•) No

## Profiling

Profiling is any form of automated processing of personal data to evaluate certain personal aspects relating to a person.
*For example analysing or predicting aspects concerning that person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements*

Will your research include creating detailed profiles of individuals which may have a long lasting or significant impact on those individuals?

- ○ Yes
- ◉ No

### Decisions Against Individuals

Will the project result in you or others making decisions, or taking action against individuals in ways which can have a significant impact on them?

- ○ Yes
- ◉ No

### Legality of Processing Section Statement

The following questions will be used to ensure that research projects are undertaken legally by determining who is responsible for deciding the purposes and means of data processing, and the legal basis for doing so.

### Data Controller

Specify the data controller(s) for personal data processed in the course of the project - this is normally the University, unless the data is processed under contract to or in partnership with another organisation. Select at least one of the following:

- ☑ Heriot-Watt University
- ☐ Other

### Legal Basis for Processing

State the legal basis for processing the personal data obtained in the course of the project. (Refer to ⓘ to determine which of the following to select).

The data subject has given consent to the processing of his or her personal data

## Processing Outside of the EEA?

**The following countries are currently in the European Economic Area (EEA):**

Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the UK

**The European Commission has designated the following countries are providing an adequate level of protection for privacy:**

Andorra, Argentina, Canada (commercial organisations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland, Uruguay and the United States of America (limited to the companies operating under Privacy Shield framework)

Will the project involve transfers of data outside of the above listed countries, to organisations that are not members of Heriot-Watt University Group?

- ○ Yes
- ◉ No

## Further Questions Required?

**Screening Process Results**

The responses you have provided to questions in the previous sections indicate that the research project is either a medium or low risk project.

If your project is potentially high risk then you will need to answer more questions called a Data Protection Impact Assessment.

Please confirm whether you agree that the project is a low or medium risk project, or if you disagree and you believe the project is potentially high risk:

- ◉ Agree: This is either a low or medium risk project
- ○ Disagree: This is a potentially high risk project

## Information Security Requirements

The following questions will be used to determine the security measures that will be applied to protect data from accidental or deliberate unauthorised disclosure, loss, or alteration.

## Information Flows

Describe the information flows in the project:

Would you like to attach a flow diagram to describe the information flows?

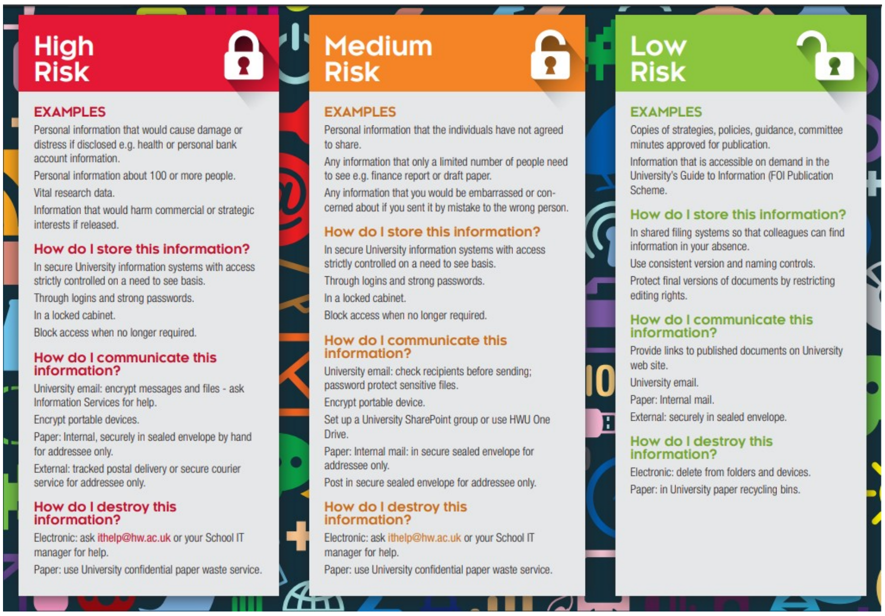
- Yes
- No

## Access Control and Training

Who will require access to the data, including Heriot-Watt staff and other partners, in order to complete work required for the research project? List any individuals and staff groups.

## Heriot-Watt Information Security Standards

The following information is provided in Heriot-Watt Information Security guidance. It specifies the standard of controls that should be put in place to protect high, medium, and low risk information. Please read through this information carefully:




In particular:

- Data should be stored in secure Heriot-Watt University Information Systems (e.g. *Heriot-Watt University Office 365 account, Heriot-Watt Home or Shared Drive, Heriot-Watt One Drive*)
- Use strong password protection
- Portable devices should be encrypted
- Paper documents with confidential information should be disposed of using University confidential waste service
- Check email recipients before sending.

Will you adhere to all the above standards for storing, communicating and destroying information?

- (x) Yes
- ( ) No

### External Third Parties

Will any data or information be sent outside of university managed IT systems (*Heriot-Watt University Office 365 account, Heriot-Watt Home or Shared Drive, Heriot-Watt One Drive etc.*)?

For example if the data is collected, stored or processed by a third party e.g. a cloud data storage company, an online survey software provider (such as Qualtrics), an application provider, market research company translator or transcriber?

- ( ) Yes – Data is stored or processed outside of Heriot-Watt IT Systems
- (x) No – Data is stored and processed within Heriot-Watt IT Systems

### Retention of Personally Identifiable Data

How long do you intend to retain personally identifiable data, *for example participant's names and contact details, or the participants' unique identifier*?

- (x) Unless explicit consent is provided for a data subject to be named in the project outputs, personally identifiable data will only be kept for as long as it is necessary to keep the data in order to verify the integrity of the project's methods and to ensure the validity of the outputs.
- ( ) Other

### Other Ethics

Are there any ethical issues that you have identified which have not yet been addressed in this application?

- ( ) Yes
- (x) No

### Other Documents

Are there any other documents which may support the ethics application which have not yet been uploaded?

- ○ Yes
- ◉ No

## Applicant

Please confirm that the details completed in this form are accurate and a true reflection of the intended research project, and that if the details of the research project change significantly that you will seek additional ethical approval:

- ☑ Confirm

Applicant Signature:

## Supervisor

I (the supervisor) am satisfied that the researcher has properly considered the ethical implications of the intended research project and has taken appropriate action.

Supervisor
Signature:

## How to Submit an Application

When your form is complete and ready to be submitted for review by your school's research ethics committee, click the **SUBMIT** button on the left hand panel of your screen