Improving Acceptance Rate for Hotel Partners' Recommendations: Generative AI personalization, contextualization and explainability

Florian Hatt*
fhatt@expediagroup.com
Expedia Group
Geneva, Switzerland

Stefano Casasso
scasasso@expediagroup.com
Expedia Group
Geneva, Switzerland

Kirubhakaran Krishnanathan
kkrishnanathan@expediagroup.com
Expedia Group
London, UK

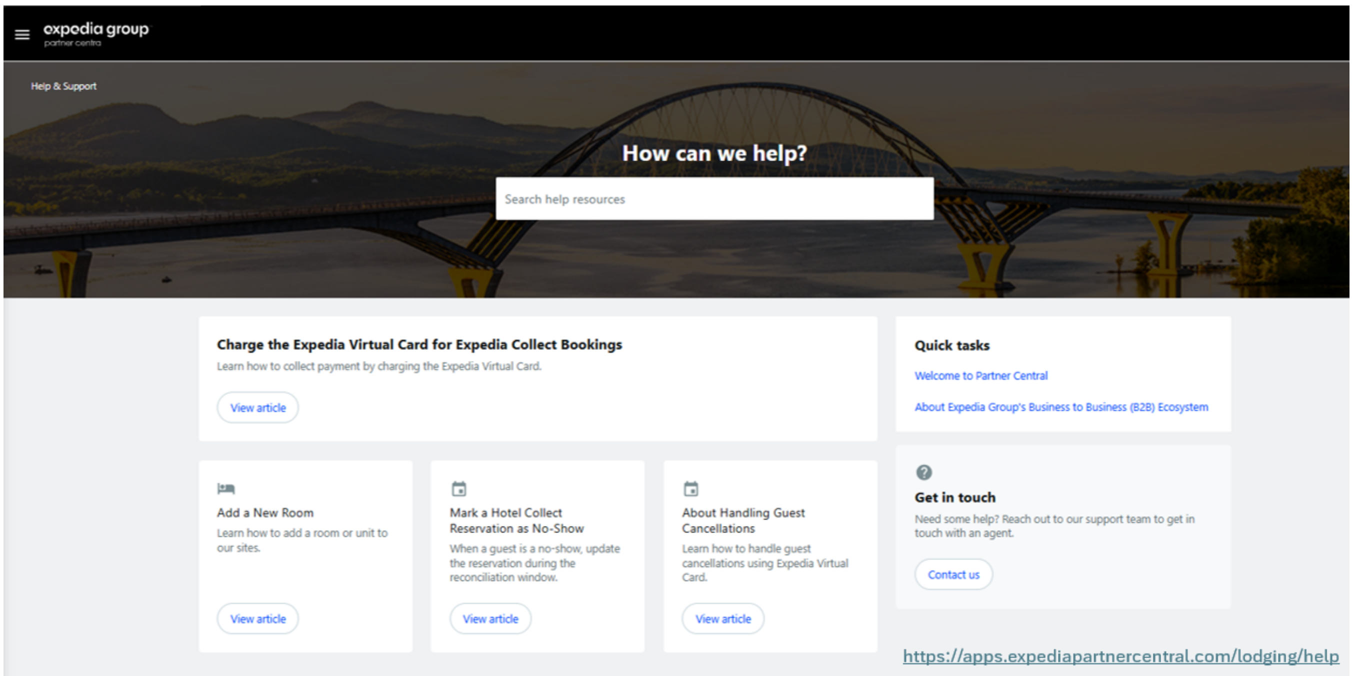

Figure 1: Expedia Partner Central - Help and Support

ABSTRACT

The online travel sector increasingly relies on data-driven decision-making to optimize business outcomes for hotel partners. Expedia's current recommendation system, based on causal inference models, estimates the impact of potential interventions but often lacks clarity and contextualization, leading to low adoption rates among hotel partners. This paper proposes integrating a generative AI layer with Expedia's causal recommender system to transform statistical recommendations into personalized, context-rich narratives. The hybrid system leverages large language models (LLMs) to enhance trust and adoption by providing customized explanations that align with individual business goals. The paper details the technical foundations, architecture, and application of this system, and proposes an application which will allow to test its business impact.

CCS CONCEPTS

• **Computing methodologies** → **Causal reasoning and diagnostics**; • **Information systems** → **Recommender systems**.

KEYWORDS

Recommender System, Causal Inference, Generative AI, Large Language Model, Explainability

ACM Reference Format:
Florian Hatt, Stefano Casasso, and Kirubhakaran Krishnanathan. 2025. Improving Acceptance Rate for Hotel Partners' Recommendations: Generative AI personalization, contextualization and explainability. In *Proceedings of*

Conference KDD 2025, Toronto,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

August 2025 (Conference KDD 2025). ACM, New York, NY, USA, 6 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the dynamic landscape of online travel platforms, Expedia stands at the forefront, facilitating connections between millions of travelers and millions of accommodation providers. Central to this ecosystem are hotel partners who rely on Expedia's tools for pricing, visibility, and demand forecasting. To optimize business outcomes such as occupancy rates, revenue per available room (RevPAR), and average daily rates (ADR), Expedia employs sophisticated recommendation systems. However, a persistent challenge remains: hotel partners often hesitate to act on algorithmically generated suggestions due to a lack of clarity, limited contextualization, and misalignment with individual business goals.

Traditional recommender systems predominantly utilize correlation-based models, which, while effective in certain domains, may not capture the causal relationships essential for strategic decision-making in the hospitality industry. Recognizing this limitation, researchers have advocated for the integration of causal inference into recommender systems to better estimate the impact of specific actions on outcomes. For instance, Gao et al. (2022) emphasize that "the real world is driven by causality, not just correlation, and correlation does not imply causation" [3] highlighting the necessity of causal models in capturing true effect relationships. Despite the enhanced predictive capabilities of causal models, their adoption among hotel partners is hindered by issues of interpretability and trust. This phenomenon, known as algorithm aversion, describes the tendency of individuals to distrust or reject algorithmic recommendations, especially when they lack transparency or fail to provide understandable justifications [4].

To bridge this trust gap, the integration of generative AI, particularly large language models (LLMs) like ChatGPT, offers a promising solution. These models can generate coherent, context-rich narratives that translate complex statistical outputs into human-understandable insights. Remountakis et al. (2023) explore this potential, stating that "by incorporating persuasive techniques, such as social proof, scarcity, and personalization, recommender systems can effectively influence user decision-making and encourage desired actions" [6]. As highlighted by Deldjoo et al. (2024) "the introduction of large language models (LLMs) such as ChatGPT and Gemini have showcased remarkable emergent capabilities, including reasoning, in-context few-shot learning, and access to extensive open-world information within their pre-trained parameters. Because of their broad generalist abilities, these pretrained generative models have opened up an exciting new research space for a wide variety of recommendation applications" [2].

In the context of Expedia, integrating a generative AI layer atop the causal inference-based recommender system aims to address the interpretability challenge. By precomputing features and generating personalized, contextualized insights, the system can provide hotel partners with clear explanations of recommended actions, thereby aligning suggestions with their business objectives and fostering greater trust in the platform's guidance.

This paper delves into the architecture of this hybrid system, examining how the synergy between causal inference and generative AI can enhance recommendation effectiveness. We present a detailed case study within the Expedia platform, demonstrating the practical applications and benefits of this approach. Furthermore, we discuss the broader implications for the hospitality industry and propose avenues for future research to refine and expand upon this integrated model.

2 LITERATURE REVIEW

2.1 Causal Inference in Recommender Systems

Traditional recommender systems primarily rely on correlation-based methods, such as collaborative filtering and matrix factorization, to predict user preferences. While effective in many scenarios, these approaches often fall short in capturing the underlying causal relationships that drive user behaviors. This limitation can lead to issues like popularity bias and exposure bias, where certain items are recommended not because they are inherently preferred by users, but because they have been historically popular or frequently exposed.

To address these challenges, researchers have increasingly turned to causal inference techniques to enhance recommender systems. Causal inference aims to estimate the effect of an intervention (e.g., recommending a particular item) on an outcome (e.g., user engagement), accounting for confounding factors that may influence both the intervention and the outcome. This approach enables more accurate predictions of how users would respond to recommendations, leading to more effective and personalized suggestions.

Various methods have been proposed to incorporate causal inference into recommender systems. These include propensity score matching, instrumental variable techniques, and counterfactual modeling. For instance, Luo et al. (2023) discuss the application of the potential outcome framework and structural causal models to address biases and improve the generalization of recommendations.

Moreover, Zhu et al. (2023) explore strategies for bias mitigation, explanation, and generalization in causal recommender systems. They highlight the importance of modeling the causal effect of item features on user feedback to address issues like confounding and selection bias.

2.2 Generative AI in Recommender Systems

A systematic literature review by researchers in 2024 analyzed 52 studies on generative AI-based recommender systems, revealing that such models often outperform traditional AI techniques in terms of accuracy and personalization. The review notes that "generative AI models, such as GANs and VAEs, have been widely used in recommender systems and they perform better than traditional AI techniques." [5]

Furthermore, Deldjoo et al. (2024) discuss the integration of large language models (LLMs) and multimodal generative models in recommender systems. They highlight the potential of these models to process and generate diverse types of content, including text, images, and videos, thereby enabling more comprehensive and context-aware recommendations [2]. The application of generative AI in recommender systems also extends to the generation of natural language explanations for recommendations, enhancing user trust and engagement . By providing users with understandable and personalized rationales for recommendations, generative models can improve the transparency and persuasiveness of recommender systems, ultimately increasing the probability of the users' to accept this recommendation.

2.3 Explainable AI in Recommender Systems

As recommender systems become increasingly integral to various digital platforms, the demand for transparency and interpretability in their decision-making processes has grown. Explainable Artificial Intelligence (XAI) addresses this need by providing insights into how and why specific recommendations are made, thereby enhancing user trust, satisfaction, and system effectiveness.

The opacity of complex models, often referred to as "black-box" systems, can lead to user skepticism and reduced trust in recommendations. Zhang and Chen (2020) emphasize that explainable recommendations help improve transparency, persuasiveness, effectiveness, trustworthiness, and user satisfaction. They note that "explainable recommendation helps to improve the transparency, persuasiveness, effectiveness, trustworthiness, and satisfaction of recommendation systems" [8].

Assessing the quality and effectiveness of explanations is crucial for the development of explainable recommender systems. Chen et al. (2022) provide a comprehensive survey on the evaluation of explainable recommendations, categorizing evaluation methods into user-centric and system-centric approaches. They highlight that "explainable recommendation has shown its great advantages for improving recommendation persuasiveness, user satisfaction, system transparency, among others". User-centric evaluations focus on how explanations affect user perceptions, trust, and decision-making, often employing user studies and surveys. System-centric evaluations, on the other hand, assess the technical aspects of explanations, such as fidelity, completeness, and computational efficiency. [1].

Traditional XAI techniques like SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-Agnostic Explanations), and counterfactual explanations have been instrumental in providing transparency to complex machine learning models. These methods offer insights into model predictions by attributing importance to input features or by illustrating how slight changes in inputs can alter outputs. However, their applicability in real-world recommender systems, especially in domains requiring nuanced, domain-specific narratives, has limitations [7].

To address the shortcomings of traditional XAI methods, integrating Large Language Models (LLMs) like GPT-4 into recommender systems offers a promising avenue. LLMs can generate coherent, context-rich narratives that translate complex statistical outputs into human-understandable insights. By leveraging both the statistical outputs of models and historical user data, LLMs can craft personalized explanations that align with users' expectations and domain-specific knowledge.

In summary, while traditional XAI techniques provide foundational transparency to recommender systems, their integration with LLMs can bridge the gap between technical explanations and user-friendly narratives. This synergy enhances the interpretability of recommendations, fosters user trust, and ultimately leads to more effective and persuasive recommender systems.

We are proposing an application and architecture to go further in this direction, using Expedia Hotels' Recommendations System.

3 APPLICATION TO EXPEDIA HOTELS' RECOMMENDATIONS WITH GENERATIVE INSIGHTS

3.1 Expedia Hotels Recommendations System

Expedia's partner ecosystem includes tens of thousands of independent hotels globally, many of which rely on the platform's insights to manage pricing, inventory, and visibility. While the existing causal inference-based recommendation system excels in computing the expected impact of potential interventions—such as adjusting weekend rates or joining a marketing campaign—adoption of these recommendations by partners remains inconsistent.

Expedia Hotels Recommendations System is a hybrid model using both causal inference uplift and collaborative filtering:

- **causal inference uplift**: Used to provide hotels with their next best action based on the return on investment of this action for themselves.
- **collaborative filtering**: Used to also integrate the learning from what similar hotels usually adopt as products

With this hybrid approach, Expedia Hotels Supply Recommendations ensure that the recommended action will provide business value to the hotels but at same time that it is also an action often adopted by similar properties.

However, in this hybrid system, there is not an easy way to provide hotels with explainability for this recommendation given the complexity of the approach, neither to show how this recommendation would help them to achieve their specific business goals.

3.2 Personalized and contextualized hotels recommendations

One of the challenges for hotels' supply recommendations to be adopted in self-service by partners are based on the following limitations:

- Lack of interpretability and narrative clarity.
- Recommendations presented without sufficient contextual framing.
- Difficulty mapping model's output to actionable partner goals

For those reasons, hotels partners can often refuse to take a recommended action despite causal uplift model has evaluated that it would be beneficial for them or despite other competitors similar hotels have adopted it.

To overcome these limitations, we propose augmenting the existing hybrid recommendation system with a generative AI layer that leverages large language models (LLMs). This layer is designed to generate natural language narratives that explain the rationale behind each recommended action in a personalized and context-sensitive manner.

4 ARCHITECTURE OF THE HYBRID GENERATIVE AI RECOMMENDATION SYSTEM

The integration of an LLM-based narrative generator on top of the causal and collaborative filtering backbone allows Expedia to bridge the interpretability gap. This approach leverages structured data (e.g., predicted uplift, feature contributions, local demand signals, and historical performance) to generate explanations that are human-readable, persuasive, and tailored to the partner's specific context.

The generative layer has three primary objectives:

- **Personalization**: Adapt the explanation language and content to reflect each hotel's size, business goals (e.g., maximizing revenue vs. increasing occupancy), and strategic behavior.
- **Contextualization**: Reference relevant external and internal signals—like competitor pricing, local events, or recent booking patterns—that justify the timing and relevance of the recommendation.
- **Explainability**: Clearly outline the causal and collaborative logic behind the recommendation in accessible language, often referencing similar past actions that were successful.

For example, rather than simply suggesting a hotel increase its weekend rate by 10%, the generative model might produce: *"Your forecasted occupancy for the upcoming weekend is 94%, which is higher than your local competitors (average 86%). Last year, during similar demand peaks, a 10% rate increase yielded a 13% revenue uplift without reducing occupancy. We recommend a similar pricing adjustment this week."*

4.1 Data used for the Generative AI layer

An important part of our Generative AI application is to source the information to use from multiple environments:

- **All interactions with our suppliers**: In this dimension, we use all the historic touch-points with this partner, from clicks on our supplier's website, emails sent and read, notes from our sales teams in Salesforce and even the transcripts of the actual pitches made.
- **All information about our supply products**: We have all our knowledge articles centralized that gather information about any supply product, how it works, its benefits and even the explanations on how to enable it.
- **All bookings data about this partners and its competitors**: In order to have relevant insights for a specific partner, we also use all their bookings data and can compare their performance on any specific segment of customers versus last year or a set of similar properties.
- **All data about eligible recommendations**: Finally, since the end goal is to support our recommendations, we plug to all the data on the eligible recommendations to this partner, and the optimal parameters for each, as well as some generic data on the uplift impact estimations of those products.

We will now describe the architecture we have used to achieve (a) personalized, (b) contextualized and (c) explainable insights to support our suppliers' recommendations.

4.2 Generative AI design for our supply insights

Maybe the first question would be: Why not just plug all those different data sources directly to a LLM with a proper prompt?

We identified multiple challenges with just integrating the data into our LLM prompts or even using a simple RAG architecture, the main challenges were about the following elements:

- Extracting from a very large amount of data, in particular the transcripts, the specific information related specifically to a supply products that were discussed with the partners.
- Similarly, computing and identifying from timeseries data the specific patterns that are worth highlighting in the supply insights because they diverge from a regular/usual situation.
- Finally and most importantly the need from the system to have very strong guardrails avoiding any risk of hallucination to compare and estimate accuracy between entry data and LLM generated output.

Therefore, we had to build a complete design that was first extracting specific data from all the data sources described in section 4.1 using more deterministic ML models, in particular:

- **NLP topics extraction** from transcripts or sales teams notes
- **Anomaly detections** systems based on time-series
- **Causal Inference** precomputed impact models

With those more classical and deterministic ML models we can fully control the information sent to the LLM step - as well as increase the time to process by running those steps in batch process and only having the LLM step as on-demand. This way, we can also entirely assess the accuracy of the LLM generated output in comparison with the pre-computed information that were sent to the LLM.

5 PROPOSED APPLICATIONS AND TESTING FRAMEWORK

5.1 Proposed design for the generative AI supply insights application

With all the previously discussed data sources we needed this application to connect to, as well as the requirements to enable guardrails system allowing us to fully control the elements exposed in the LLM outputs, we proposed and implemented a very specific design based on the diagram illustrated in Fig.2.

5.2 Guardrails system and accuracy evaluation

As explained in section 4.2, the guardrails and accuracy evaluation is a very crucial part of this Generative AI application to guarantee the personalized and contextualized supply insight is (a) accurate with the data from our different sources and (b) aligned with the supply recommendations available.

For this task, we propose to use 4 main evaluation systems in order to estimate the similarity between the inputs data provided to the LLM and the generated output:

- BERT-based Contextual Similarity
- ROUGE-N scores.
- Matching of metrics in output
- Other LLM evaluation of inputs/outputs similarity

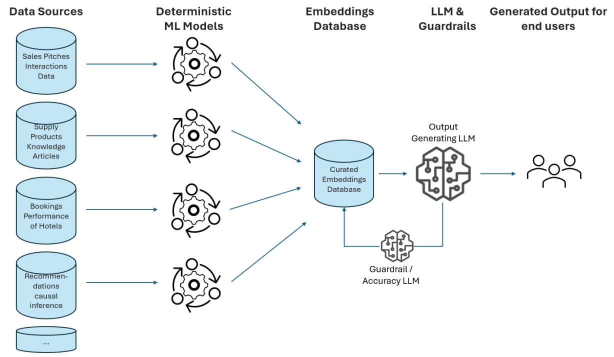

Figure 2: Proposed design for the generative AI supply insights application

5.3 Illsutration of the Generated Insights for a Recommendation

Our proposed design allows us to start from a specific supply recommendation and enrich it with (a) contextualized, (b) personalized and (c) explainable insights supporting the recommended action.

In Fig.3 and Fig.4, we provide examples of this enrichment process for 2 different use cases:

- Recommendation an Advertising Program
- Recommendation for a Last Minute Discount

Data / Information	AI Generated Insights contextualized & personalized to the Partner
No Insight (current default solution)	We recommend you to load a last-minute Discount of 10% to attract more travelers
Adding Performance Data Insights	We observe that **you're performing below similar properties in your area on the segment of the last-minute searches** and have a risk of unsold rooms. We recommend to load a last-minute Discount of 10% to get your faire share on this segment.
Adding insights from historic interactions	**As discussed last week, your objective is to fill all your rooms in this low season period**. The last-minute bookings can help you with this objective and it's a segment where you're under-performing compared to similar properties. Load a 10% last-minute Discount to reach your goal.

Figure 3: Illustration of Generated Insights for Last Minute Discount Recommendation

5.4 Proposed testing framework

With this Generative AI design, Expedia can test its business impact by looking at the Acceptance Rate of our recommended actions in 2 main consecutive use cases:

- **Human in the loop**: we propose an initial testing framework with our sales teams in an A/B test design with a beta-tester group having access to the generated supply insights (and a feedback loop to improve based on the provided feedback) - we can then observe both the human satisfaction with the generated insights, as well as the impact on the acceptance rates of sales teams using the tool against the other ones.
- **Self-Served recommendations**: in a second phase, we propose another A/B test approach without Human In the Loop control, by displaying directly to our supply partners the recommendation with and without the generated supply insights. Similarly we can observe on a larger scale the direct impact of our insights on the Acceptance Rate of the recommendations on this direct self-served channel.

Data / Information	AI Generated Insights contextualized & personalized to the Partner
No Insight (current default solution)	We recommend you to participate to EG Advertising Program for September month with 5% bidding to optimize your revenue.
Adding Performance Data Insights	We observe recently **a decrease of your performance that is mainly related to lower page views during September period**. To solve for this challenge, we would recommend to participate to EG Advertising Program for September with 5% bidding.
Adding insights from supply product knowledge articles	We observe recently a decrease of your performance that is mainly related to lower page views during September period. **This pay-per-stay model enhances your visibility in the highly competitive Honolulu market without compromising your ADR, making it an efficient solution**. We can help you to enable an EG Advertising Program for September with 5% bidding.

Figure 4: Illustration of Generated Insights for LAdvertising Campaigns

6 DISCUSSION AND NEXT STEPS

The integration of a generative AI layer into Expedia's hotel partner recommendation system marks a significant step toward addressing long-standing challenges in explainability, contextualization, and partner trust. While causal inference and collaborative filtering provide robust, data-driven decision frameworks, their complexity often limits practical adoption. By incorporating a narrative layer powered by Large Language Models (LLMs), we bridge the gap between statistical rigor and business interpretability.

6.1 Key Benefits Observed

Our implementation demonstrates several tangible advantages:

- **Improved interpretability:** Partners are better able to understand not just what the recommendation is, but why it is being made, increasing their likelihood to act.
- **Greater contextual alignment:** Recommendations are framed using partner-specific data, recent behavior, and comparative market context, resulting in more relevant and trusted guidance.
- **Scalability and operational impact:** The architecture enables personalized communication across thousands of partners, reducing dependency on human intervention while improving consistency and quality.

These benefits align with recent academic findings on the use of LLMs for explainable recommendation generation [6–8], and suggest that natural-language rationales are not only desirable but operationally transformative.

6.2 Remaining Challenges

Despite promising outcomes, several challenges remain:

- **LLM factual reliability:** Ensuring generated narratives remain consistent with precomputed data is a non-trivial task. While our guardrails system mitigates hallucination risk, continuous monitoring and benchmarking are required.
- **Real-time performance:** Batch processing of precomputed insights supports latency-sensitive applications, but full real-time responsiveness (e.g., live scenario simulation) may require architectural refinements.
- **Partner heterogeneity:** Hotel partners differ in goals, tech-savviness, and decision styles. Building adaptive explanation strategies that vary tone, detail, and complexity remains an open research area.
- **Data Sparsity and Cold Start Problem:** for some of the hotels in Expedia platform, for instance newly acquired properties, a limited number of data points for personalization is available (touchpoints, performance insights, etc.)

6.3 Next Steps and Research Opportunities

Based on our experience and feedback, the following directions are planned:

- **Interactive explanations:** Enable hotel partners to query the insight layer dynamically, asking follow-up questions such as "Why not a different action?" or "What happened last time I did this?"
- **Reinforcement learning from feedback:** Incorporate partner engagement metrics (e.g., whether the recommendation was followed) and satisfaction surveys to fine-tune future narratives.
- **Localized and multilingual expansion:** Extend support to global partners by developing LLM prompts in multiple languages and regional business contexts, leveraging translation-aware generation.
- **Integration with account managers:** Embed insights directly into sales enablement tools, using generated narratives to assist in partner conversations and onboarding materials.
- **Visual+Verbal interfaces:** Combine generated text with data visualizations (e.g., uplift graphs, competitor benchmarks) to improve clarity and engagement through multimodal presentation.

6.4 Toward a Recommender System Copilot

Ultimately, we envision the generative insight layer as the foundation for a broader partner-facing decision support system—an AI copilot that not only recommends actions, but explains, persuades, and learns from its users over time. This model blends predictive intelligence with human-centered design, pushing the frontier of enterprise recommender systems.

REFERENCES

[1] Chen, X., Zhang, Y., Wen, J.-R. 2022. Measuring Why in Recommender Systems: A Comprehensive Survey on the Evaluation of Explainable Recommendation. *arXiv:2202.06466* (2022).

[2] Deldjoo, Y., He, Z., McAuley, J., Korikov, A., Sanner, S., Ramisa, A., Vidal, R., Sathiamoorthy, M., Kasirzadeh, A., Milano, S. 2024. A Review of Modern Recommender Systems Using Generative Models (Gen-RecSys). *ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (2024).

[3] Gao, C., et al. 2022. Causal Inference in Recommender Systems: A Survey and Future Direction. *arXiv:2404.00579v2* (2022).

[4] Jussupow, E., Benbasat, I., Heinzl, A. 2020. Why Are We Averse Towards Algorithms? A Comprehensive Literature Review on Algorithm Aversion. *ECIS 2020 Proceedings* (2020).

[5] Matthew Ojo Ayemowa, Roliana Ibrahim,Muhammad Murad Khan. 2024. Analysis of Recommender System Using Generative Artificial Intelligence: A Systematic Literature Review. *DOI: 10.1109* (2024).

[6] Remountakis, M., Kotis, K., Kourtzis, B., Tsekouras, G. E. . 2023. ChatGPT and Persuasive Technologies for the Management and Delivery of Personalized Recommendations in Hotel Hospitality. *arxiv.org/pdf/2307.14298* (2023).

[7] Salih, A., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Menegaz, G., Lekadir, K. 2023. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. *arXiv:2305.02012* (2023).

[8] Zhang, Y., Chen, X. . 2020. Explainable Recommendation: A Survey and New Perspectives. *arXiv:1804.11192* (2020).