Asia Pacific Journal of Information System and Digital Transformation

2024 , Vol 1 No 01, Research Article

Factors Affecting User Satisfaction and Loyalty of the MITA Chatbot Bank Mandiri

Riswan Septriayadi Sianturi
Department of Information Systems, Universitas Brawijaya, Indonesia
rsianturi@ub.ac.id

M. Alfian Noor
Department of Information Systems, Universitas Brawijaya, Indonesia
muhammadalfn@student.ub.ac.id

Issa Arwani
Department of Information Systems, Universitas Brawijaya, Indonesia
issa.arwani@ub.ac.id

Deden Witarsyah
Department of Information Systems, Universitas Brawijaya, Indonesia
Universiti Tun Hussein Onn Malaysia, Malaysia
deden@uthm.edu.my

Abstract: The adoption of artificial intelligence (AI)-based technologies in customer service has increased substantially in recent years, particularly through the deployment of chatbots capable of simulating human conversation to support and guide users. Within the banking sector, chatbots have become essential tools for improving service accessibility, operational efficiency, and user engagement. Bank Mandiri introduced MITA (Mandiri Intelligent Assistant) as a non-transactional AI-based chatbot designed to provide customers with product and service information quickly and conveniently. However, a decline in MITA's performance ranking, coupled with a rise in user complaints, suggests the presence of service quality issues that may affect both user satisfaction and loyalty. To address these concerns, this study investigates the determinants of user satisfaction and loyalty toward MITA by employing the AI Bot Service Quality model, which includes Core AI Bot Service Quality, AI Bot Service Recovery Quality, and AI Bot Conversational Quality. The model is further extended to incorporate key Artificial Intelligence Features, including Trendiness, Visual Attractiveness, Problem-Solving Ability, and Communication Quality, to provide a more comprehensive understanding of the factors influencing user perceptions of chatbot-based banking services. A quantitative approach was used, with data collected through online questionnaires distributed to 126 active MITA users located on Java Island. Structural Equation Modeling using Partial Least Squares (SEM-PLS) was applied to test the proposed hypotheses. The results reveal that Core AI Bot Service Quality, Problem-Solving Ability, and Communication Quality significantly enhance user satisfaction. Moreover, user satisfaction emerges as a strong predictor of user loyalty, while Core AI Bot Service Quality indirectly influences loyalty through satisfaction. Overall, the findings highlight the importance of enhancing the fundamental capabilities of AI chatbots, specifically their accuracy in problem resolution and the clarity of their communication, to enhance user experience and foster long-term loyalty toward AI-enabled digital banking services.

Keywords: **Chatbot; Artificial Intelligence Features; Service Quality; Satisfaction; Loyalty**

1. INTRODUCTION

Artificial Intelligence (AI) is a broad branch of computer science that encompasses various disciplines, focusing primarily on developing systems capable of performing tasks that typically require human intelligence (Malabadi et al., 2023). Unlike humans, who have performance limitations, AI can continuously process data without interruption. To accelerate the adoption of this technology, the Indonesian government has formulated the National Artificial Intelligence Strategy 2020-2045 (BPPT, 2020).

In the banking sector, chatbots play a crucial role in improving customer engagement and operational efficiency. Banks utilize AI-powered chatbots to provide 24/7 support, automate repetitive tasks, and deliver personalized experiences. According to Agnihotri & Bhattacharya, (2024)Chatbots have become essential tools for financial institutions seeking to enhance user satisfaction and streamline digital service delivery.

Bank Mandiri, one of Indonesia's largest state-owned banks, launched MITA (Mandiri Intelligent Assistant) as an AI-driven non-transactional chatbot designed to offer product information and general service guidance. However, despite its promising implementation, MITA's performance has shown a downward trend. User complaints on social media about delayed responses, irrelevant answers, and limited conversational depth suggest a decline in service quality that could affect user satisfaction and loyalty.

Service quality has long been recognized as a key factor in determining customer satisfaction and loyalty. A study conducted by Dantsoho and Ringim revealed that the quality of artificial intelligence in financial services enhances customer satisfaction, which in turn encourages customers to continue using the service (Aliyu Dantsoho et al., 2021). Many theoretical models have been proposed and developed to explain the factors of electronic service quality that drive customer satisfaction and loyalty. Parasuraman introduced the E-S-QUAL and E-RecS-QUAL frameworks in 2005 as extensions of the SERVQUAL model (Parasuraman et al., 2005). Researchers later modified these frameworks to measure service quality in AI-driven chatbot environments. For example, the AI Bot Service Quality Model incorporates AI bot conversational quality as a variable related to chatbot interaction quality (Hsu and Lin, 2023). This model identifies three main dimensions: Core Service Quality, Service Recovery Quality, and Conversational Quality, which collectively shape users' perceptions of chatbot performance. While the model effectively captures the technical and interaction aspects of service quality, it does not fully address the evolving characteristics of modern AI-based systems, particularly those related to user experience and intelligent system behavior.

Artificial intelligence in the banking sector has been examined primarily in the context of chatbot adoption. (Kwangsawad and Jattamart, 2022). Omoge et al., (2022) highlighted that AI has enhanced customer satisfaction within the banking industry. Hsu & Lin, (2023) has provided insights into chatbot service quality but has not addressed the AI-related characteristics that may influence customer satisfaction. Therefore, this study identifies four AI characteristic factors that affect customer satisfaction with banking chatbot services. The AI Bot Service Quality model is integrated with these AI characteristic factors to provide a deeper understanding of how chatbot service quality, AI characteristics, and other factors interact with each other.

2. Research Model and Hypotheses

2.1. Research Model

The AI Bot Service Quality model, developed by Hsu & Lin, (2023) extends the concept of electronic service quality into AI-driven environments. It comprises three main dimensions: Core AI Bot Service Quality, AI Bot Service Recovery Quality, and AI Bot Conversational Quality. Core Service Quality refers to the chatbot's ability to deliver accurate, relevant, and timely information. Service Recovery Quality measures how effectively the chatbot handles errors or failures in conversation, while Conversational Quality reflects how naturally and emotionally engaging the chatbot is in communicating with users.

Figure 1 illustrates the model used in the present study, highlighting that chatbot interactions differ from traditional web services due to their dynamic and conversational nature. User satisfaction is influenced not only by the chatbot's functional quality but also by its ability to understand intent, respond contextually, and express empathy. This model offers a structured framework for evaluating the technical and emotional aspects of AI-based service systems.

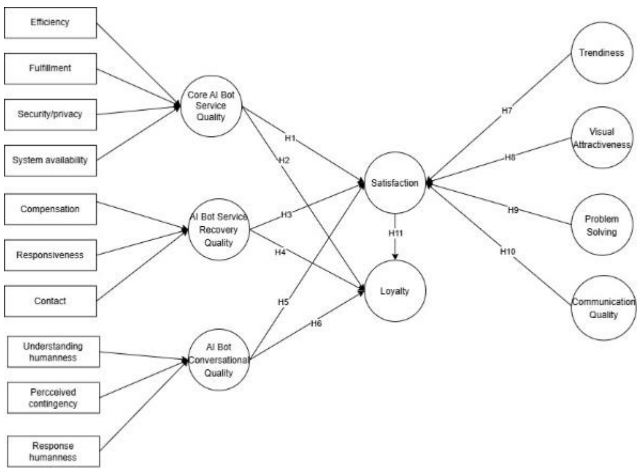

Fig. 1 Research Model

2.2. Hypothesis

a) Core AI Bot Service Quality and AI Bot Service Recovery Quality

With the rapid growth of electronic services, previous studies have demonstrated that chatbot service quality has a significant impact on user satisfaction (Chen et al., 2023). Furthermore, Zhou et al., (2021) found that the quality of electronic banking services influences customer loyalty. Zehir and Narcıkara (2016) found that both service quality and service recovery quality have a significant impact on customers' intentions to remain loyal. Similarly, Hsu and Lin (2023) discovered that the quality of

AI chatbot service recovery plays a crucial role in enhancing customer satisfaction. The information and theories discussed above serve as the foundation for the following hypotheses:

- H1: Core AI Bot Service Quality positively influences Satisfaction
- H2: Core AI Bot Service Quality positively influences Loyalty
- H3: AI Bot Service Recovery Quality positively influences Satisfaction
- H4: AI Bot Service Recovery Quality positively influences Loyalty

b) AI Bot Conversational Quality

Human-like chatbots have been shown to impact customer satisfaction with luxury brands that utilize chatbots positively. (Chung et al., 2020). Hsiao and Chen (2022) found that the human-like impression of chatbots and the quality of service in food-ordering chatbots are key factors that shape customer satisfaction. Maroufkhani et al. (2022) discovered that users' perceptions of human-like voice assistants significantly influence their intention to continue using them. Hsu & Lin (2023) stated that the quality of chatbot conversations, characterized by the ability to understand user messages, maintain conversational flow, and provide human-like responses, has the most significant effect on user satisfaction but does not directly impact loyalty. Jenneboer et al. (2022) also found that human-like chatbots increase customer satisfaction and trust, which in turn leads to consumer loyalty. Therefore, the following hypothesis is proposed:

- H5: Core AI Bot Service Quality positively influences Satisfaction
- H6: Core AI Bot Service Quality positively influences Loyalty

c) Artificial Intelligence Features

In the context of AI-powered digital banking services, trendiness refers to the service provider's ability to deliver up-to-date information about digital banking features, including the latest trends and developments in the banking industry. According to Mi Alnaser et al., (2023), trendiness significantly contributes to enhancing user satisfaction. It is considered one of the key features of AI-based banking services due to its ability to align with the lifestyles of users who seek continuously updated and modern services. This is because users perceive that they are engaging with services that stay relevant to current developments and support their digital identity. Chung et al., (2020) also indicated that AI-driven digital banking services can meet the needs and modern lifestyles of customers, ultimately leading to higher user satisfaction. In addition to trendiness, visual appeal can also evoke feelings of satisfaction and reduce users' intentions to switch to other services. (Kuo, 2020). An aesthetically pleasing interface design can enhance comfort during interactions and create a positive user experience. (Bhandari et al., 2019). In a study conducted by Candiwan and Annikmah, (2024), visual attractiveness was found to have a positive effect on user satisfaction, as an appealing appearance can influence emotions and create a professional impression that strengthens trust in the service. Accordingly, we hypothesize:

- H7: Trendiness positively influences Satisfaction
- H8: Visual Attractiveness positively influences Satisfaction

Furthermore, Mi Alnaser et al., (2023) emphasized that problem-solving ability is a crucial element in technology, including artificial intelligence in digital banking, as it enhances customer satisfaction by continuously providing practical solutions. The ability of service agents to handle customer issues efficiently and sincerely ultimately influences customers' perceptions of their satisfaction levels (Chung et al., 2020). A chatbot's capability to resolve problems serves as a key indicator of system performance. Candiwan and Annikmah (2024) revealed that the problem-solving aspect of AI-based digital banking services significantly affects user satisfaction, as it reflects the technological sophistication in delivering fast and accurate solutions. Another essential feature of AI-based digital banking services is the quality of communication. Communication quality refers to the level of effectiveness in providing information from service professionals to clients, which includes the accuracy of information, trust, response speed, and time efficiency (Wang and Li, 2012). Elements such as accuracy, credibility, and competence in conveying information are considered key indicators in evaluating the effectiveness of communication (Chung et al., 2020). Accuracy pertains to how well the information provided aligns with customer needs and context, while credibility reflects the degree to which customers trust the source of the information. Competence relates to the ability of agents or systems to respond to inquiries in a relevant and convincing manner. Candiwan and Annikmah (2024) emphasized that communication quality significantly contributes to user satisfaction in digital banking, as effective communication builds trust and ensures that users receive the information they need accurately and promptly. The hypothesis proposed is:

- H9: Problem Solving positively influences Satisfaction
- H10: Communication Quality positively influences Satisfaction

d) Satisfaction and Loyalty

User satisfaction and loyalty have always been the primary goals of all companies. Therefore, this study proposes that satisfaction has a positive effect on loyalty in chatbot usage. According to Kotler and Armstrong (Kotler and Armstrong, 2018), satisfaction is defined as a person's feeling of pleasure or disappointment that arises after comparing the perceived performance of a product or service with their expectations. Previous studies have found that user satisfaction has a significant influence on loyalty and the intention to continue using the service in the future (So et al., 2021). The hypothesis proposed is:

- H11: Satisfaction positively influences Loyalty

3. RESEARCH METHODOLOGY

3.1. Quantitative Approach

This study employed a quantitative approach to investigate the relationships between AI Bot Service Quality dimensions, Artificial Intelligence Features, Satisfaction, and loyalty among MITA chatbot users. The target population included all active users of the Mandiri Intelligent Assistant (MITA) across Java Island, Indonesia. According to data from the Indonesia Deposit Insurance Corporation (LPS) in 2023, the provinces on the island of Java recorded the highest number of bank customers in the country. This dominance indicates a high level of banking service utilization in Java, making respondents from this region relevant for representing customer behavior and perceptions toward the diverse range of digital banking services available in the area. The questionnaire was adapted from previously validated scales developed by Hsu and Lin, (2023) and Mi Alnaser et al., (2023), which were then modified to fit the context of banking chatbots. Data were collected through a questionnaire using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Data were collected through an online questionnaire distributed via Google Forms between March and May 2024. Respondents were selected using purposive sampling, with inclusion criteria requiring that each respondent had used the MITA chatbot at least once.

The data were processed and analyzed using SmartPLS 4.1.1.2 software with the PLS-SEM (Partial Least Squares–Structural Equation Modeling) method. PLS is a composite-based SEM approach that emphasizes predictive accuracy when evaluating composite models with causal-explanatory structures (Hair et al., 2019). SEM is a multivariate data analysis technique commonly applied in marketing research because it enables the testing of theoretical models supported by linear and additive causal relationships, facilitating the examination of variable interconnections and the prioritization of resources (Kwong and Wong, 2015). Compared to other statistical modeling approaches, PLS-SEM offers more precise estimations, particularly when dealing with small sample sizes (Russo and Stol, 2022). Before conducting the survey, a pilot study was carried out using an online questionnaire involving 35 respondents to identify potential weaknesses in the measurement instrument (Ghozali, 2006) and to ensure that all survey items were clear and easily understood. A total of 126 valid responses were obtained after data screening.

3.2. Demographics of Respondents

The majority of respondents were digitally literate individuals within the productive age group, who represent the primary target market for AI-based services, such as MITA, in Indonesia's banking sector (Table 1). Their high exposure to digital platforms makes them suitable respondents for evaluating the performance and user experience of AI-driven chatbots.

TABLE 1. DEMOGRAPHICS STATISTICS

Chatacteristics	N
Gender	
Male	46
Female	80
Age	
17-20	11
21-25	71
26-30	35
>30	9
Domicile	
DKI Jakarta	43
East Java	35
West Java	22
Central Java	18
DI Yogyakarta	4
Banten	4
Job	
Student	58
Private Employee	55
Enterpreneur	5
Civil Servant	5
Housewife	2
Trader	1
Duration of Use	
Less than 3 months	26
3-6 months	36
6-12 months	25
1-2 years	21

Chatacteristics	N
More than 2 years	18

4. DATA ANALYSIS AND RESULTS

4.1. Measurement Model Evaluation

The PLS-SEM model and hypotheses were analyzed in two stages. The first stage involved conducting a PLS analysis of the item-construct relationships as part of the instrument validation process. The second stage tested the hypothesized relationships through structural model analysis and bootstrap resampling (Henseler et al., 2009). In addition, the measurement model evaluation requirements had to be met to ensure both convergent validity (for first-order and second-order constructs) and discriminant validity. A two-stage embedded approach was employed to analyze the higher-order constructs as an alternative to the repeated indicator approach (Sarstedt et al., 2019).

a) First Order Construct

The measurement model was evaluated by examining the reliability, convergent validity, and discriminant validity of both first-order and second-order constructs. Convergent validity is achieved when the factor loadings are above 0.70, and the Average Variance Extracted (AVE) for each construct is greater than 0.50 (Hair et al., 2017), while reliability is confirmed when the values of Cronbach's Alpha exceed 0.60, and Composite Reliability (CR) exceeds 0.70. Since the factor loading value of EFI3 was below 0.70, it was excluded from the analysis. The data presented in Table 2 indicate that all measurement indicators met the required criteria.

TABLE 2. CONVERGENT VALIDITY FIRST ORDER

Indicator	Loading Factor	AVE	CR
EFI1	0,768	0,536	0,830
EFI2	0,700		
EFI4	0,778		
FUL1	0,807	0,634	0,874
FUL2	0,788		
FUL3	0,818		
FUL4	0,771		
SEC1	0,811	0,618	0,866
SEC2	0,805		
SEC3	0,714		
SEC4	0,810		
STA1	0,918	0,800	0,889
STA2	0,870		
COM1	0,886	0,739	0,850
COM2	0,833		
RES1	0,871	0,696	0,873
RES2	0,797		
RES3	0,833		
CON1	0,867	0,780	0,876
CON2	0,898		
USH1	0,934	0,843	0,942
USH2	0,920		
USH3	0,900		
PCC1	0,839	0,677	0,893
PCC2	0,810		
PCC3	0,804		
PCC4	0,838		
RPH1	0,906	0,837	0,939
RPH2	0,924		
RPH3	0,916		

Thus, the first-order construct measurement model was found to be convergently valid and internally consistent. Discriminant validity was confirmed using the Heterotrait-Monotrait Ratio (HTMT) method, which calculates the average correlation between indicators of different constructs. As shown in Table 3, good discriminant validity is indicated by all values being below the conservative threshold of 0.85 (Henseler et al., 2015).

TABLE 3. DISCRIMINANT VALIDITY (HTMT) FIRST ORDER

	COM	CON	EFI	FUL	PCC	RES	RPH	SEC	STA	USH

	COM	CON	EFI	FUL	PCC	RES	RPH	SEC	STA	USH
COM										
CON	0.621									
EFI	0.593	0.279								
FUL	0.694	0.648	0.712							
PCC	0.655	0.593	0.426	0.685						
RES	0.504	0.596	0.541	0.778	0.508					
RPH	0.124	0.111	0.244	0.371	0.349	0.426				
SEC	0.706	0.518	0.617	0.442	0.78	0.486	0.143			
STA	0.388	0.702	0.526	0.744	0.483	0.625	0.312	0.239		
USH	0.227	0.088	0.343	0.412	0.468	0.35	0.711	0.291	0.353	

b) Second Order Construct

The two-stage approach utilized latent variable scores from the first-order model to construct the second-order measurement model. The second-order measurement model was also evaluated by examining convergent validity, reliability, and discriminant validity. Convergent validity was assessed using factor loadings and AVE values, with most indicators meeting the acceptable thresholds of greater than 0.7 for factor loadings (except for the security/privacy indicator) and 0.5 for AVE (Hair et al., 2017). Indicators that did not meet the factor loading threshold were eliminated. In terms of reliability, the variable visual attractiveness did not meet the minimum Cronbach's alpha requirement of 0.6 and was therefore removed from the model. The data presented in Table 4 demonstrate that all measurement indicators satisfied the established evaluation criteria.

TABLE 4. CONVERGENT VALIDITY SECOND ORDER

Indicator	Loading Factor	AVE	Cronbach's Alpha	CR
EFI	0,769	0,665	0,752	0,856
FUL	0,892			
STA	0,734			
COM	0,715	0,600	0,680	0,817
RES	0,859			
CON	0,743			
USH	0,837	0,638	0,718	0,841
PCC	0,757			
RPH	0,801			
TRE1	0,888	0,741	0,652	0,851
TRE2	0,833			
PRS1	0,753	0,774	0,860	0,910
PRS2	0,939			
PRS3	0,934			
CQU1	0,796	0,660	0,744	0,854
CQU2	0,829			
CQU3	0,812			
SAT1	0,901	0,777	0,857	0,913
SAT2	0,902			
SAT3	0,840			
LOY1	0,867	0,771	0,852	0,910
LOY2	0,899			
LOY3	0,869			

Finally, the HTMT validity results indicated an overlap between AI Bot Service Recovery Quality and Communication Quality, with a correlation value of 0.942. Similarly, a significant correlation of 0.953 was found between AI Bot Service Recovery Quality and Core AI Bot Service Quality. These associations suggest that respondents may perceive these constructs as closely related, which could compromise the model's clarity and validity. To improve discriminant validity and reduce the HTMT values below the 0.90 threshold, it is recommended to remove items showing a high average correlation with items from other constructs (Hair et al., 2017). Therefore, the items' compensation and system availability were removed, as they exhibited the highest average correlations. Table 5 shows that discriminant validity was achieved, with all values falling below the conservative threshold of 0.85 after removing highly correlated items.

TABLE 5. DISCRIMINANT VALIDITY (HTMT) SECOND ORDER

	AI-BCQ	AI-BSRQ	CQU	CAI-BSQ	LOY	PRS	SAT	TRE
AI-BCQ								
AI-BSRQ	0.625							

	AI-BCQ	AI-BSRQ	CQU	CAI-BSQ	LOY	PRS	SAT	TRE
CQU	0.688	0.844						
CAI-BSQ	0.688	0.882	0.881					
LOY	0.622	0.616	0.704	0.758				
PRS	0.608	0.374	0.375	0.291	0.364			
SAT	0.674	0.581	0.708	0.758	0.88	0.52		
TRE	0.473	0.692	0.641	0.43	0.274	0.104	0.342	

5. Data analysis and results

Path coefficients were tested using SmartPLS 4 to evaluate the structural model. A hypothesis was considered statistically significant at the 5% significance level when the p-value was less than 0.05, and the t-value exceeded 1.96 (Monecke and Leisch, 2012). The model's effect size was measured using the coefficient of determination (R^2). According to the guidelines proposed by (Henseler et al., 2009), R^2 values of 0.75, 0.50, and 0.25 are classified as substantial, moderate, and weak. The R^2 value for loyalty was 0.607, with an adjusted R^2 of 0.594, indicating a moderate explanatory power. This means that the independent variables in the model accounted for 59.4% of the variance in loyalty. Meanwhile, the satisfaction variable had an R^2 value of 0.529 and an adjusted R^2 of 0.505, suggesting that 50.5% of the variance in satisfaction was explained by the independent variables, which also falls within the moderate category.

Additionally, hypothesis testing was conducted by calculating the estimated path coefficients, t-statistics, and p-values obtained through the bootstrap procedure. The bootstrap method applied in this analysis involved 5,000 samples with a 5% significance level. A hypothesis was considered supported when the t-statistic exceeded 1.96, and the p-value was less than 0.05 (Monecke and Leisch, 2012). Table 6 shows that out of the ten hypotheses tested, four hypotheses were accepted (H1, H9, H10, H11), while six hypotheses were rejected (H2, H3, H4, H5, H6, H7).

TABLE 6. PATH COEFFICIENT

Hypothesis	Direction	Path Coefficient	T Statistics	P Value	Decision
H1	CAI-BSQ → SAT	0.370	3.664	0.000	Accepted
H2	CAI-BSQ → LOY	0.161	1.295	0.098	Not Accepted
H3	AI-BSRQ → SAT	-0.013	0.127	0.449	Not Accepted
H4	AI-BSRQ → LOY	0.060	0.587	0.279	Not Accepted
H5	AI-BCQ → SAT	0.130	1.591	0.056	Not Accepted
H6	AI-BCQ → LOY	0.063	0.901	0.184	Not Accepted
H7	TRE → SAT	-0.005	0.052	0.479	Not Accepted
H9	PRS → SAT	0.270	3.782	0.000	Accepted
H10	CQU → SAT	0.185	1.668	0.048	Accepted
H11	SAT → LOY	0.596	6.100	0.000	Accepted

a) Core AI Bot Service Quality on Satisfaction

The analysis results for Hypothesis H1 demonstrate a positive and statistically significant relationship between Core AI Bot Service Quality and user satisfaction ($\beta = 0.370$, $t = 3.664$, $p < 0.001$). Accordingly, H1 is supported. This suggests that higher levels of fundamental service quality provided by the chatbot are associated with increased user satisfaction during interactions. This result supports the findings of Chen et al., (2023), which highlights that the fundamental quality of AI chatbot services is a key factor in shaping satisfaction. In the context of the MITA chatbot, this finding reinforces that users are more satisfied when the chatbot can efficiently perform essential tasks such as providing banking information promptly, maintaining user privacy and security while accessing the chatbot, and operating smoothly without technical issues.

b) Core AI Bot Service Quality on Loyalty

The analysis results for Hypothesis H2 indicate a positive but statistically insignificant relationship between Core AI Bot Service Quality and user loyalty ($\beta = 0.161$, $t = 1.295$, $p = 0.098$). Therefore, H2 is not supported. This finding suggests that although the chatbot's basic service quality may be adequate, it does not have a significant impact on users' loyalty. This finding contrasts with (Hsu and Lin, 2023) who found that the fundamental service quality of chatbots has a significant impact on user loyalty. The discrepancy may be explained by the demographic data, which show that most respondents had been using the chatbot for less than one year, suggesting that their limited experience may not yet be sufficient to develop loyalty. In the context of financial services, which are often perceived as critical and sensitive, users tend to associate their loyalty more with trust in the bank's overall reputation rather than with the chatbot's performance itself.

c) AI Bot Service Recovery Quality on Satisfaction

The analysis results for Hypothesis H3 indicate a non-significant relationship between Recovery Quality and user satisfaction ($\beta = –0.013$, $t = 0.127$, $p = 0.449$). Accordingly, H3 is not supported. This suggests that the recovery actions provided during chatbot failures do not significantly impact user satisfaction. This finding differs from the study of Hsu and Lin (2023), which found a significant effect of service recovery quality on satisfaction. In the context of MITA, this result can be explained by respondents' answers, which indicate that the basic services of the chatbot have already been adequately fulfilled. In other words, MITA chatbot users rarely experience system failures that require recovery processes, so service recovery quality is not a dominant factor in shaping their satisfaction. The effectiveness of service recovery largely depends on users' prior experiences with service failures. If failures occur infrequently or go unnoticed, the quality of recovery tends not to significantly influence satisfaction perceptions.

d) AI Bot Service Recovery Quality on Loyalty

The analysis results for Hypothesis H4 indicate that the relationship between Service Recovery Quality and continuance intention is not statistically significant ($\beta = 0.060$, $t = 0.587$, $p = 0.279$). Accordingly, H4 is not supported. This suggests that the service recovery efforts provided by the chatbot do not meaningfully influence users' intentions to continue using the system in the future. This finding is consistent with the results of Hsu and Lin (2023) who also found that service recovery quality does not have a direct effect on chatbot user loyalty. However, it differs from the findings of [15], which showed that service recovery quality affects user loyalty in online retail sites. This rejection may be due to differences in service characteristics. In chatbots, the recovery process is automatic and impersonal, so users do not experience the emotional engagement that can foster loyalty.

e) AI Bot Conversational Quality on Satisfaction

The analysis results for Hypothesis H5 indicate a non-significant relationship between Conversational Quality and user satisfaction ($\beta = 0.130$, $t = 1.591$, $p = 0.056$). Accordingly, H5 is not supported. This suggests that the naturalness of the chatbot's conversational responses does not directly contribute to increasing user satisfaction. This result contradicts the findings of Hsu and Lin (2023), which revealed that chatbot conversational quality has a significant effect on user satisfaction. In the context of the MITA chatbot of Bank Mandiri, although the chatbot can provide responses that resemble human conversation, this is not sufficient to increase satisfaction if the responses do not fully address users' financial needs.

f) AI Bot Conversational Quality on Loyalty

The analysis results for Hypothesis H6 indicate a non-significant relationship between Natural Conversational Quality and user loyalty ($\beta = 0.063$, $t = 0.901$, $p = 0.479$). Accordingly, H6 is not supported. This suggests that the chatbot's natural conversational ability does not significantly impact users' loyalty. This result contrasts with the findings of Jenneboer et al., (2022), who reported that chatbots with human-like characteristics encourage user loyalty in commerce-based chatbot interactions. Conversely, it aligns with Hsu and Lin (2023), who found that conversational quality does not have a direct impact on chatbot user loyalty. In the context of the MITA chatbot of Bank Mandiri, this can be explained by the nature of digital banking services, where users primarily view chatbots as functional tools for completing service needs rather than as interactive partners that foster emotional connections. Consequently, even though the chatbot demonstrates human-like conversational abilities, this is not sufficient to strengthen users' intention to continue using the service.

g) Trendiness on Satisfaction

The analysis results for Hypothesis H7 indicate a non-significant relationship between Information Currency and user satisfaction ($\beta = -0.005$, $t = 0.052$, $p = 0.479$). Accordingly, H7 is not supported. This suggests that although users receive up-to-date information about Bank Mandiri's services through the MITA chatbot, the timeliness of information alone does not substantially contribute to their satisfaction. This finding contradicts Candiwan and Annikmah, (2024), who found that trendiness affects user satisfaction in AI-based digital banking. The distinct characteristics of chatbot use in banking can explain the difference. The MITA chatbot is primarily utilized for quick service needs, such as checking balances or transaction information, meaning that user satisfaction is evaluated based on functional aspects rather than lifestyle-oriented factors.

h) Visual Attractiveness on Satisfaction

The analysis results for hypothesis H8 indicate that the construct of the Trendiness variable is unreliable, with a Cronbach's Alpha value of 0.582, which falls below the minimum threshold of 0.70. This suggests that the indicators used to measure this variable lack sufficient internal consistency to represent the visual attractiveness construct as a whole. After conducting the outer model assessment, it was found that the visual interface aspect of MITA is not strong enough to be consistently measured within this research model. This is likely because the MITA chatbot operates through a text-based platform (WhatsApp), rather than a standalone visual interface such as a fully interactive application. Therefore, the measurement instrument used in this study may not adequately capture users' perceptions of MITA's visual appeal.

i) Visual Attractiveness on Satisfaction

The analysis results for Hypothesis H9 demonstrate a positive and statistically significant relationship between the chatbot's problem-solving capability and user satisfaction with MITA ($\beta = 0.270$, $t = 3.782$, $p < 0.001$). Accordingly, H9 is supported. This suggests that the higher levels of problem-solving effectiveness provided by the chatbot are associated with increased user satisfaction. This result is consistent with the findings of Mi Alnaser et al., (2023) and Candiwan and Annikmah (2024), which highlights that problem-solving ability has a significant effect on user satisfaction in AI digital banking contexts in Indonesia. The findings further emphasize that MITA's effectiveness in providing accurate and prompt solutions to customer issues plays a crucial role in enhancing user satisfaction. Users tend to feel satisfied when the chatbot can resolve their questions or complaints without human intervention, particularly in situations that require immediate responses.

j) Communication Quality on Satisfaction

The analysis results for Hypothesis H10 demonstrate a positive and statistically significant relationship between Communication Quality and user satisfaction ($\beta = 0.185$, $t = 1.668$, $p = 0.048$). Accordingly, H10 is supported. This suggests that higher levels of communication quality provided by the MITA chatbot contribute to increased user satisfaction during interactions. This result is consistent with the findings Candiwan and Annikmah (2024) and Mi Alnaser et al., (2023) , which revealed that high-quality communication in AI-based banking services significantly enhances user satisfaction. Users feel satisfied when the chatbot delivers relevant information in a concise and easily understandable manner. Clear and professional communication within the MITA chatbot reflects its credibility as a digital financial service platform.

k) Satisfaction on Loyalty

The analysis results for Hypothesis H11 demonstrate a positive and statistically significant relationship between user satisfaction and user loyalty ($\beta = 0.596$, $t = 6.100$, $p < 0.001$). Accordingly, H11 is supported. This indicates that users who experience higher satisfaction when interacting with the MITA chatbot are more likely to develop stronger loyalty toward its continued use. This result aligns with the study of Hsu and Lin (2023), which identified satisfaction as a key determinant of user loyalty toward chatbots. In the context of MITA, this finding emphasizes that user satisfaction plays a crucial role in encouraging long-term engagement. When users are satisfied with the chatbot's service quality and AI-based features, that satisfaction tends to motivate them to continue using the chatbot and recommend it to others.

6. CONCLUSION

This study demonstrates that Core AI Bot Service Quality, Problem-Solving Capability, and Communication Quality significantly enhance user satisfaction with the MITA chatbot, and that satisfaction subsequently exerts a strong positive influence on user loyalty. These findings underscore that loyalty in chatbot-based digital banking is primarily constructed through positive service experiences—particularly when the chatbot performs its essential functions reliably, provides accurate and timely solutions to user problems, and communicates information clearly and professionally. Despite these contributions, several limitations should be noted. The study's sample size was relatively limited, which may restrict the generalizability of the results. Future research should consider expanding the sample across diverse demographic segments and longer usage durations to capture more representative user experiences. Moreover, incorporating additional variables such as e-trust, perceived value, technology readiness, or corporate reputation may provide deeper insights into the mechanisms that shape satisfaction and loyalty in AI-driven banking services. Longitudinal or comparative studies across different chatbot platforms could further enrich understanding of how chatbot quality dimensions influence user perceptions and behavioral intentions over time.

REFERENCES

[1] Agnihotri A, Bhattacharya S. Chatbots' effectiveness in service recovery. Int J Inf Manage 2024;76. https://doi.org/10.1016/j.ijinfomgt.2023.102679.
[2] Aliyu Dantsoho M, Jinjiri Ringim K, Maitama Kura K. The Relationship between Artificial Intelligence (AI) Quality, Customer Preference, Satisfaction and Continuous Usage Intention of e-Banking Services. Indonesian Business Review, 4 (1), 24–43 2021.
[3] Bhandari U, Chang K, Neben T. Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. Information & Management 2019;56:85–93. https://doi.org/10.1016/j.im.2018.07.003.
[4] BPPT. National Strategy for Artificial Intelligence 2020-2045. 2020.
[5] Candiwan C, Annikmah RR. Exploring the Impact of Artificial Intelligence on User Satisfaction and Acceptance in Digital Banking Services in Indonesia. 2024 IEEE 30th International Conference on Telecommunications, ICT 2024, Institute of Electrical and Electronics Engineers Inc.; 2024. https://doi.org/10.1109/ICT62760.2024.10606022.
[6] Chen Q, Lu Y, Gong Y, Xiong J. Can AI chatbots help retain customers? Impact of AI service quality on customer loyalty. Internet Research 2023;33:2205–43. https://doi.org/10.1108/INTR-09-2021-0686.
[7] Chung M, Ko E, Joung H, Kim SJ. Chatbot e-service and customer satisfaction regarding luxury brands. J Bus Res 2020;117:587–95. https://doi.org/10.1016/j.jbusres.2018.10.004.
[8] Ghozali Imam. Aplikasi analisis multivariate dengan program SPSS. Badan Penerbit Universitas Diponegoro; 2006.
[9] Hair JF, Hult GTM, Ringle CM, Sarstedt M, Thiele KO. Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. J Acad Mark Sci 2017;45:616–32. https://doi.org/10.1007/s11747-017-0517-x.

[10] Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. European Business Review 2019;31:2–24. https://doi.org/10.1108/EBR-11-2018-0203.

[11] Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2015;43:115–35. https://doi.org/10.1007/s11747-014-0403-8.

[12] Henseler J, Ringle CM, Sinkovics RR. The use of partial least squares path modeling in international marketing, 2009, p. 277–319. https://doi.org/10.1108/S1474-7979(2009)0000020014.

[13] Hsiao K-L, Chen C-C. What drives continuance intention to use a food-ordering chatbot? An examination of trust and satisfaction. Library Hi Tech 2022;40:929–46. https://doi.org/10.1108/LHT-08-2021-0274.

[14] Hsu CL, Lin JCC. Understanding the user satisfaction and loyalty of customer service chatbots. Journal of Retailing and Consumer Services 2023;71. https://doi.org/10.1016/j.jretconser.2022.103211.

[15] Jenneboer L, Herrando C, Constantinides E. The Impact of Chatbots on Customer Loyalty: A Systematic Literature Review. Journal of Theoretical and Applied Electronic Commerce Research 2022;17:212–29. https://doi.org/10.3390/jtaer17010011.

[16] Kotler P, Armstrong G. Principles of Marketing. 17th Edition. Pearson Education; 2018.

[17] Kuo R-Z. Why do people switch mobile payment service platforms? An empirical study in Taiwan. Technol Soc 2020;62:101312. https://doi.org/10.1016/j.techsoc.2020.101312.

[18] Kwangsawad A, Jattamart A. Overcoming customer innovation resistance to the sustainable adoption of chatbot services: A community-enterprise perspective in Thailand. Journal of Innovation and Knowledge 2022;7. https://doi.org/10.1016/j.jik.2022.100211.

[19] Kwong K, Wong K. Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS. 2015.

[20] Malabadi RB, Kolkar KP, Chalannavar RK, Mudigoudra BS, Abdi G, Baijnath H. International Journal Of Research And Innovation In Applied Science (IJRIAS) Applications of Artificial Intelligence (AI) in Cannabis Industries: In Vitro Plant Tissue Culture 2023. https://doi.org/10.51584/IJRIAS.

[21] Maroufkhani P, Asadi S, Ghobakhloo M, Jannesari MT, Ismail WKW. How do interactive voice assistants build brands' loyalty? Technol Forecast Soc Change 2022;183:121870. https://doi.org/10.1016/j.techfore.2022.121870.

[22] Mi Alnaser F, Rahi S, Alghizzawi M, Ngah AH. Does artificial intelligence (AI) boost digital banking user satisfaction? Integration of expectation confirmation model and antecedents of artificial intelligence enabled digital banking. Heliyon 2023;9. https://doi.org/10.1016/j.heliyon.2023.e18930.

[23] Monecke A, Leisch F. semPLS : Structural Equation Modeling Using Partial Least Squares. J Stat Softw 2012;48. https://doi.org/10.18637/jss.v048.i03.

[24] Omoge AP, Gala P, Horky A. Disruptive technology and AI in the banking industry of an emerging market. International Journal of Bank Marketing 2022;40:1217–47. https://doi.org/10.1108/IJBM-09-2021-0403.

[25] Parasuraman A, Zeithaml VA, Malhotra A. E-S-QUAL a multiple-item scale for assessing electronic service quality. J Serv Res 2005;7:213–33. https://doi.org/10.1177/1094670504271156.

[26] Russo D, Stol K-J. PLS-SEM for Software Engineering Research. ACM Comput Surv 2022;54:1–38. https://doi.org/10.1145/3447580.

[27] Sarstedt M, Hair JF, Cheah J-H, Becker J-M, Ringle CM. How to Specify, Estimate, and Validate Higher-Order Constructs in PLS-SEM. Australasian Marketing Journal 2019;27:197–211. https://doi.org/10.1016/j.ausmj.2019.05.003.

[28] So KKF, Kim H, Oh H. What Makes Airbnb Experiences Enjoyable? The Effects of Environmental Stimuli on Perceived Enjoyment and Repurchase Intention. J Travel Res 2021;60:1018–38. https://doi.org/10.1177/0047287520921241.

[29] Wang W, Li H. Factors influencing mobile services adoption: a brand-equity perspective. Internet Research 2012;22:142–79. https://doi.org/10.1108/10662241211214548.

[30] Zehir C, Narcıkara E. E-Service Quality and E-Recovery Service Quality: Effects on Value Perceptions and Loyalty Intentions. Procedia Soc Behav Sci 2016;229:427–43. https://doi.org/10.1016/j.sbspro.2016.07.153.

[31] Zhou Q, Lim FJ, Yu H, Xu G, Ren X, Liu D, et al. A study on factors affecting service quality and loyalty intention in mobile banking. Journal of Retailing and Consumer Services 2021;60. https://doi.org/10.1016/j.jretconser.2020.102424.

DOI: https://doi.org/10.61973/apjisdt.v10124.2