Responsible AI in Community Banking: Human Oversight, Complaint Handling, and Service Accountability

Arif Budi Raharja[1], Maya Ariyanti[2]

1. Universitas Mathla'ul Anwar, Banten | budiraharja1970@gmail.com

2. Telkom University | ariyanti@telkomuniversity.ac.id

Abstract

Purpose: This conceptual study develops a framework for governing AI-assisted customer service in community banking. Approach: A purposive narrative synthesis connects service quality, operational resilience, accountability, and AI risk management literature published before November 2026. No primary observations or statistical estimates are reported. Conceptual contribution: The framework distinguishes task delegation, decision authority, customer explanation, and effective redress. It proposes that automation supports service quality when staff can identify consequential errors, suspend unsafe recommendations, and resolve complaints through an accountable process. Organisational conditions include accessible alternative channels, proportionate record keeping, and independent challenge. Implications: Banks should evaluate completed and corrected customer journeys rather than automated response volume alone. The paper offers proposed controls, indicators, and research propositions; their effectiveness requires subsequent empirical testing.

Keywords: responsible artificial intelligence; community banking; human oversight; complaints; service accountability

1. Introduction

Community banks may introduce AI into document classification, customer enquiries, complaint triage, or staff knowledge search. These uses differ substantially from autonomous lending decisions. A system that retrieves an opening-hours policy creates a different exposure from one that tells a customer why a transfer failed. Treating every application as the same technology obscures the allocation of responsibility. This article therefore examines a bounded question: how should a bank organise human oversight and redress when AI participates in a customer service journey?

The problem is especially relevant where relationship-based service remains important. Customers can depend on branch employees to interpret documents, explain unfamiliar procedures, and recognise exceptional circumstances. Automation can shorten routine interactions, but the removal of a human contact point can make an unusual problem harder to solve. The proposed framework treats service completion, understandable explanation, and correction of mistakes as complementary objectives. It does not assume that a faster initial answer produces a better final outcome.

The conceptual starting points are deliberately distinct. Parasuraman et al. (2005) provide a service-quality perspective; the Basel Committee on Banking Supervision (2021) addresses operational resilience; Tabassi (2023) provides an AI risk-management framework; and Bovens (2007) clarifies accountability relationships. Their combination motivates an organisational design argument, not a claim that these sources tested community-bank AI. The contribution is a set of conditional mechanisms and observable questions that a subsequent field study could investigate.

2. Theoretical foundations and construct boundaries

Service quality concerns what a customer can actually accomplish and how reliably the service supports that purpose. In this article, AI-assisted service quality means accurate, usable assistance within the bank's authorised service scope. It is distinct from fluency, apparent friendliness, and the number of answers generated. E-S-QUAL identifies dimensions relevant to electronic service assessment (Parasuraman et al., 2005), but its measures cannot simply be renamed and assumed valid for generative systems. New contexts require instrument development and validation.

Human oversight is defined here as an employee's practical ability to inspect a consequential output, challenge its use, and alter the resulting action. Merely placing an employee's name beside a system does not establish that ability. Oversight requires sufficient information, time, authority, and a feasible alternative procedure. A reviewer who cannot see the applicable policy or cannot pause a transaction is performing a narrower administrative task, even if the organisation calls the arrangement human review.

Accountability adds an answerability relationship between an actor and a forum capable of questioning conduct (Bovens, 2007). Applied to banking service, the question is who explains an incorrect answer, who decides the remedy, and who examines repeated failures. Operational resilience adds continuity when a supporting process fails (Basel Committee on Banking Supervision, 2021). The resulting framework joins prevention, detection, correction, and learning while preserving their different purposes. None of these constructs substitutes for the others.

3. Conceptual method and scope

The study uses purposive conceptual synthesis rather than a systematic review or an empirical evaluation. Sources were selected for their relevance to four defined problems: electronic service quality, AI risk, operational continuity, and institutional answerability. The selected references are listed at the end of the paper. There is no claim of an exhaustive database search, a screened population of publications, independent coding, interviews, a customer sample, or a statistical test.

The analytical procedure begins by defining the service boundary and the consequences of an incorrect output. It then separates tasks that a system may assist from decisions that remain assigned to authorised employees. The next step traces an error from generation through delivery, customer interpretation, complaint submission, and resolution. Finally, the analysis identifies conditions that could interrupt correction, such as unavailable records or a reviewer lacking authority. These steps organise the reasoning; they are not a description of an observed bank implementation.

Three assumptions delimit the argument. First, the bank remains responsible for its customer-facing services even where an external supplier provides the model. Second, lawful access to relevant records and an appropriate basis for processing customer information must exist before a workflow is implemented. Third, customers need a feasible route for assistance without having to understand model architecture. The recommendations are organisational proposals. They require local legal, technical, and operational assessment and should not be interpreted as proof of compliance with any particular regulation.

4. A service-accountability framework

The proposed causal sequence begins with the allocation of a bounded task. A bank specifies which information sources the application may use, which actions it may recommend, and which circumstances require escalation. Staff then compare outputs with an authorised policy or record. Customer-facing explanations identify the next available action without implying that a probabilistic answer is an official determination. Complaints supply evidence about failures that escaped these controls.

The framework predicts a conditional relationship: automation is more likely to improve usable service when the bank also strengthens correction capacity. Without that capacity, repeated fast answers can increase the burden on the customer. This is an analytical proposition rather than an estimated effect. It also allows an alternative outcome: a well-governed tool can still add little value where the underlying procedure is unnecessarily complex. Process simplification and system governance therefore need

separate evaluation.

Table 1 translates the argument into control questions. The functions are complementary. A bank should not use a strong policy document as evidence that frontline review is effective, or use low complaint counts as evidence that the complaints route is accessible. An explicit exception process is particularly important because users experiencing the greatest difficulty may also face the highest cost of describing the problem.

Table 1. Proposed framework

Control stage	Accountability question
Task boundary	Which outputs are authorised and which require escalation?
Human challenge	Can staff access evidence and change the action?
Customer explanation	Does the customer understand the next available step?
Redress	Who investigates, remedies, and records closure?
Learning	Who examines recurrence and changes the process?

5. Designing a bounded operating model

A practical design begins with a task inventory. Management records the service purpose, information sensitivity, permitted output, accountable owner, and fallback route. Low-consequence tasks can be assessed separately from those involving financial loss, access restriction, or a potentially misleading statement of rights. Tabassi (2023) provides a general risk-management orientation; the specific inventory proposed here is an application developed for this paper. It is not a certification instrument or a validated scoring model.

The second design decision concerns escalation. An employee should receive the original enquiry, the material system output, the relevant evidence, and a clear statement of what remains unresolved. Requiring the customer to repeat the entire history creates a new failure at the handover point. A bank can define escalation triggers around uncertainty, conflicting records, customer vulnerability, and repeated unsuccessful attempts.

Thresholds must be chosen and tested locally; this paper does not prescribe numerical limits.

The third decision concerns continuity. If the application becomes unavailable or its outputs become unreliable, staff require a documented alternative. The alternative may be a controlled policy repository, a telephone queue, or a branch appointment. Its capacity matters: a fallback that cannot accept the redirected workload offers only nominal protection. The bank should test a plausible interruption scenario and examine whether customers retain access to essential assistance. Such a test would generate implementation evidence that this conceptual study itself does not supply.

6. Measurement and implementation evidence

Evaluation should distinguish system behaviour from customer outcomes. Technical output accuracy can be assessed against an appropriately constructed reference set. Service completion concerns whether the customer's valid objective was achieved. Redress concerns whether an identified error was corrected, explained, and closed without unreasonable repetition. A single average cannot adequately represent these dimensions. Table 2 offers candidate indicators, each requiring a defined denominator, observation window, and responsible owner before use.

Complaint data are informative but selected. A low rate can reflect good service, a difficult reporting channel, fear of consequences, or a customer's decision to abandon the process. Complaint volume should therefore be interpreted alongside channel availability and unresolved-contact records. E-S-QUAL and the service-recovery literature motivate attention to the user journey (Parasuraman et al., 2005), but an AI-specific instrument must still establish content validity, reliability, and comparability across channels.

Implementation should begin with a restricted service scope and documented review of exceptions. A before-and-after comparison could be misleading if enquiry complexity or staffing changes at the same time. A stronger future study would record case mix, channel, severity, and relevant organisational changes, and would predefine its primary

outcomes. Customer privacy should be protected throughout sampling and analysis. The purpose of measurement is to discover failure and guide correction, rather than to produce a favourable adoption narrative.

Table 2. Candidate indicators and interpretation safeguards

Candidate indicator	Interpretation safeguard
Completed eligible journeys / eligible requests	Separate complexity and channel before comparison.
Corrected material errors / identified material errors	Report unresolved cases and observation window.
Escalations with adequate evidence / escalations	Assess relevance, not document volume alone.
Repeat contacts for the same unresolved problem	Link cases carefully and protect personal information.

7. Discussion and testable propositions

Proposition 1 states that the association between AI assistance and service completion depends on effective employee authority to challenge outputs. This proposition is stronger than the claim that human presence is sufficient. It would be contradicted if nominally supervised systems performed equally well regardless of employees' access to evidence, time, and decision authority. A future study should distinguish these components instead of treating oversight as a binary label.

Proposition 2 states that accessible redress reduces the persistence of consequential service errors, conditional on the bank's capacity to investigate and implement remedies. This does not predict that accessible complaint channels immediately reduce recorded complaints. Better access may initially increase reporting. Proposition 3 states that operational continuity moderates the relationship between technology dependence and service disruption. A tested alternative channel should matter most when the primary tool fails, rather than only during normal operation.

These propositions leave room for competing explanations. Better-managed banks may both adopt stronger controls and deliver better service, creating selection effects. More complex cases may be routed to human staff, making raw comparisons unfair. Customers may prefer different channels for reasons unrelated to service quality. Empirical research should address these possibilities through transparent sampling, severity adjustment, and careful temporal design. The framework's value is to make the assumptions explicit, not to declare any bank or technology successful in advance.

8. Managerial implications and limitations

The managerial priority is an accountable service journey. A useful operating question is whether a named employee can reconstruct and correct a consequential interaction using the evidence available at the time. Reconstruction should be proportionate: recording every possible detail can introduce avoidable privacy and security exposure. The bank needs a retention rationale, controlled access, and a clear link between records and the purpose of investigation. These are design requirements to assess locally, not a statement that more data always improve governance.

Supplier arrangements should support the same service purpose. Management can require meaningful incident reporting, an understood change process, and a way to withdraw an unreliable feature. Responsibility for a customer remedy should not disappear into a dispute about whose component produced the error. The framework also suggests that staff training should include when to refrain from using the tool. Competence includes recognising an unsupported answer and moving to an authorised alternative.

The study has substantial limitations. It contains no observed implementation, customer testimony, or causal estimate. Its purposive references cannot establish completeness of the literature. The discussion does not assess particular products, Indonesian regulatory compliance, or different model architectures. Community banks vary in staffing, infrastructure, and customer needs. Future research should examine these differences and test whether the proposed controls improve outcomes sufficiently to justify their costs. The framework should be revised when evidence contradicts its assumptions.

9. Conclusion and research agenda

AI-assisted service should be assessed as an organisational arrangement linking task boundaries, human authority, customer explanation, and correction. The central conceptual contribution is a distinction between producing a response and taking responsibility for its consequences. A bank may improve the first while weakening the second. An accountable design makes that trade-off visible and provides a route to investigate it.

The proposed framework combines several complementary perspectives without claiming that their integration has already been empirically validated. It identifies control questions and candidate measures, and it explains why complaints and continuity should be considered alongside output accuracy. Its practical recommendation is to start with a bounded service purpose, identify an accountable owner, and test whether an error can be detected and remedied in practice. Expansion should follow evidence about completed journeys and manageable failures.

Future studies can compare tasks with different consequences, examine the quality of human escalation, and evaluate customer access to redress. Researchers should report both favourable and unfavourable outcomes and distinguish observed results from design aspirations. This article offers no numerical performance promise. Its propositions are intended to support disciplined inquiry into whether a particular implementation improves service while preserving answerability to the people who depend on it.

References

Basel Committee on Banking Supervision. (2021). Principles for operational resilience. Bank for International Settlements. https://www.bis.org/bcbs/publ/d516.htm

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. European Law Journal, 13(4), 447-468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL: A multiple-item scale for assessing electronic service quality. Journal of Service Research, 7(3), 213-233. https://doi.org/10.1177/1094670504271156

Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1