SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Hybrid Ai-Assisted Agent Models FOR Insurance Operations: A Multi-Layered Architecture FOR Workflow Automation AND Human-Centered Decision Support

Maruthi Prasad Gundla

Independent Researcher. E-mail: maruthigundla@gmail.com

Abstract
Across the insurance industry, agents carry a paradox: customers need them most during claims and complex service events, yet the systems surrounding those agents offer little support in real time. Back-office platforms activate after calls end. Consumer digital portals replace agents entirely. CRM tools deliver yesterday's data but cannot synthesize what is happening during an active customer interaction. Written from the perspective of a practicing insurance systems engineer and following a design science research approach, this paper describes a four-layer architecture – Data Integration, AI Intelligence, Agent Interaction, and Customer Transparency – built to close that gap. At its center sits a claims support engine that operates during live customer calls, pulling policy terms, retrieving external incident data from telematics and weather feeds, generating a sequenced question protocol matched to the specific loss type, and delivering a complete first notice of loss package to the claims system before the conversation ends. Comparative analysis across seven capability dimensions shows no current system type meets more than three. The proposed architecture meets all seven, with projected reductions of 30–55% in quoting cycle time and gains of 35–50 percentage points in first-call FNOL documentation completeness, grounded in published carrier efficiency research. The architecture is further mapped against the NAIC model bulletin and the EU AI Act, positioning explainable hybrid routing and retained human decision authority as compliance assets rather than constraints.

Keywords: Hybrid AI, Insurance Automation, Claims Support Engine, Workflow Orchestration, Human-AI Collaboration, Agent Augmentation, Explainable AI, Design Science Research.

1. Introduction

Walk into any mid-size carrier operation and the tension becomes visible quickly. Agents are on the phone with customers who need answers – about a claim they want to open, a renewal they did not understand, a policy change mid-term. Meanwhile those same agents are tabbing between three systems, re-entering data that exists somewhere in the stack but will not surface without manual effort. The work that actually requires judgment – reading the customer, catching the inconsistency in their loss account, deciding whether to push back or let something go – gets squeezed into whatever time is left after the overhead.

Accenture (2022) put numbers to this: inadequate claims handling is projected to place roughly $170 billion in annual global premium revenue at risk by 2027 through cancellations and failed renewals. Consumer research points in the same direction. Accenture's global insurance consumer study found that 49% of consumers trust a human advisor when making an insurance claim, while only 12% trust an automated digital service and just 7% trust a chatbot (Accenture, 2021). What the data reflects is not a preference for inefficiency – it is a reasonable response to the reality that a car accident or a flooded basement is not a self-service transaction. Customers want someone on the line who knows the policy and can tell them what happens next.

Three technology categories are commonly deployed in insurance today, and each falls short in a different way. Back-office automation platforms – document classifiers, adjudication engines, fraud scoring models – activate after

the conversation has concluded. Direct-to-consumer digital channels do not involve an agent at all, which works for simple endorsements but collapses under the weight of a complex loss. CRM and quoting tools sit in front of the agent but pull from static data snapshots with no capacity to reach a weather feed, a telematics provider, or a police report database mid-call. None of these systems was designed around the agent-customer interaction as its primary unit. The architecture described in this paper is.

The paper makes four contributions. First, it specifies a four-layer reference architecture that treats the live agent-customer interaction, rather than the back-office process, as the primary unit of design. Second, it details a real-time claims support engine whose dynamic question generation capability does not exist in any current system category. Third, it evaluates the architecture through a structured comparative analysis across seven capability dimensions and a literature-grounded performance framework expressed as testable propositions. Fourth, it maps the architecture against the regulatory expectations now forming around AI in insurance, including the NAIC model bulletin and the EU AI Act. Throughout, the perspective is that of a practicing engineer: the architecture is specified at the level of systems, integrations, data flows, and latency constraints that an implementation team would build against, rather than as an abstract capability model. The remainder of the paper proceeds as follows: Section 2 positions the work in prior research, Section 3 describes the research approach, Sections 4 and 5 present the architecture and the claims support engine, Sections 6 and 7 evaluate them, Section 8 addresses regulatory, privacy, and ethical considerations, Section 9 discusses applications, and Sections 10 and 11 close with limitations and conclusions.

2. Research Background and Related Work

2.1. How Prior Research Has Framed AI-Human Task Division

Jarrahi (2018) offered one of the cleaner frameworks for thinking about AI and human professionals working in tandem. His argument was not that machines would eventually absorb most occupational tasks, but that the two modes of intelligence operate from fundamentally different strengths and that smart deployment means matching task type to the right capability. Speed across structured data favors the machine. Judgment under genuine ambiguity – where the right answer depends on context that cannot be extracted from any database – favors the person. Huang and Rust (2018) added useful granularity through four task categories: mechanical, analytical, intuitive, and empathetic. AI substitution advances most reliably through the first two and stalls on the latter two, a conclusion their subsequent work on feeling-dominant service tasks reinforced (Huang and Rust, 2021). An agent on a loss report call is cycling through all four simultaneously – retrieving data (mechanical), checking coverage applicability (analytical), reading whether the customer account feels complete (intuitive), and managing someone who is upset or frightened (empathetic). The hybrid design problem is deciding which of those four the machine should carry.

Eling et al. (2022), working from 91 academic papers and 22 industry studies, found strong and consistent evidence that AI investment generates measurable efficiency gains across the insurance value chain. They also identified a structural limitation that motivates the current work: nearly every implementation they reviewed was siloed. Fraud detection, pricing models, document intake systems – each built independently, each optimized for a narrow function, none of them coordinated into the agent-facing workflow where the customer relationship actually plays out. Duan et al. (2019) made a parallel point from an organizational decision-making lens: in environments where data quality varies and decision consequences are high, AI systems should synthesize and present rather than resolve autonomously. That principle drives the routing logic described in Section 4.

2.2. Claims Research, NLP Applications, and Customer Trust

Zhang et al. (2023) examined road traffic accident claim disputes and found that AI-generated cost estimates could accelerate resolution timelines by giving parties a shared, data-grounded starting point for negotiation. The finding matters here because it confirmed that AI augmentation in claims does not require removing the human from the loop – it requires giving the human better inputs to work with. Li et al. (2025) took a different angle, applying large language models directly to claims text. Slightly over one in five samples in their dataset contained substantive inconsistencies between the narrative and the supporting documents, gaps that early detection could convert into either a corrected submission or a flag for closer review. Their system processed each sample in under four seconds, settling the practical question about whether real-time NLP in a live call context carries prohibitive latency costs. Schrijver et al. (2024), reviewing 50 studies of machine learning in automobile insurance fraud detection, found that models perform well on constructed research datasets but that integration into real-time, agent-facing workflows remains largely unrealized – a maturity gap this architecture is designed to close. The trust element has its own set of design constraints. Owens et al. (2022) reviewed 103 papers on explainable AI in insurance and

concluded that building explanations into the decision process itself has a materially greater effect on appropriate reliance than supplying transparency after the fact as a compliance artifact. Agents who understand why a recommended escalation occurred act on the recommendation more consistently than agents shown a conclusion without a justification. Nguyen et al. (2022) added the customer side of the same coin: what customers respond to is not the sophistication of the underlying model but the quality of the communication it produces – accurate, relevant, and timely information delivered in a form a non-expert can act on. Both points shaped how the Agent Interaction and Customer Transparency layers in this architecture were designed.

2.3. The Current Wave: Agentic AI and the Regulatory Response

The period since these studies has compressed the distance between research and deployment. Industry analyses now describe agentic AI – systems that plan and execute multi-step workflows rather than answer single queries – as the dominant near-term investment theme in claims and underwriting, while cautioning that most deployments remain pilots concentrated in back-office functions (Deloitte, 2025). The regulatory environment has moved in parallel. The NAIC model bulletin on insurers' use of artificial intelligence systems, adopted by the NAIC in December 2023 and since taken up by a majority of U.S. states, requires insurers to maintain written AI governance programs, document how AI-driven decisions affecting consumers are validated, and preserve human accountability for adverse outcomes (NAIC, 2023). In the European Union, the AI Act classifies AI systems used for risk assessment and pricing in life and health insurance as high-risk, imposing transparency, human-oversight, and data-governance obligations that shape how any customer-facing decision-support system must be built (European Union, 2024). Two implications follow for the present work. The persistent concentration of AI capability in the back office confirms that the gap Eling et al. (2022) identified – siloed engines, none coordinated into the agent-facing workflow – remains open. And the direction of regulation strongly favors architectures in which the human agent retains decision authority and the system can explain itself, which is precisely the hybrid configuration developed here.

3. Research Approach

This paper follows the design science research paradigm, which treats a purposefully designed artifact – here, a reference architecture – as a legitimate research contribution when the problem is relevant, the design is rigorously grounded, and the evaluation is explicit about what it does and does not establish (Hevner et al., 2004). The work tracks the six activities of the design science research methodology: problem identification, definition of solution objectives, design and development, demonstration, evaluation, and communication (Peffers et al., 2007). Sections 1 and 2 establish the problem and its grounding in prior research. The solution objectives follow directly from that literature: real-time operation during the live interaction, multi-source data integration, explainable human-AI task routing, and customer-facing transparency. The design itself is practice-led: it consolidates integration patterns drawn from the author's engineering work with the core systems involved – policy administration, CRM, claims, and quoting platforms – and uses the published literature to validate, bound, and quantify those patterns rather than as their source. Sections 4 and 5 present the designed artifact, the application scenarios in Section 9 serve as illustrative demonstration, and Sections 6 and 7 carry the evaluation.

The evaluation uses two complementary instruments, and it is important to be precise about their status. The first is a structured comparative analysis (Section 6) in which the proposed architecture and the three system categories dominant in current insurance practice are rated across seven capability dimensions. Four of the seven dimensions operationalize the solution objectives directly (real-time agent augmentation, multi-source data integration, hybrid AI-human routing, customer transparency output); two capture the claims-specific capabilities the engine introduces (dynamic question generation, automated FNOL package assembly); and the seventh, interaction audit trail, reflects the audit requirements documented in the explainability literature (Owens et al., 2022). Ratings of Yes, Partial, or No were assigned from the documented capabilities of each category in the sources cited beneath Table 1, and the rationale for the least obvious cells is discussed in the text. The second instrument is a literature-grounded performance framework (Section 7) expressed as testable propositions with quantified expected ranges. Those ranges are drawn from published implementation research, not from a deployment of this architecture; they are claims to be tested, and Section 10 describes the production validation planned as the next phase of this work.

4. System Architecture

The architecture runs across four layers shown in Figure 1. Keeping them functionally distinct was deliberate: it allows the data retrieval logic to be updated without touching the agent dashboard, and the customer output layer to be refined for regulatory compliance without breaking anything upstream.

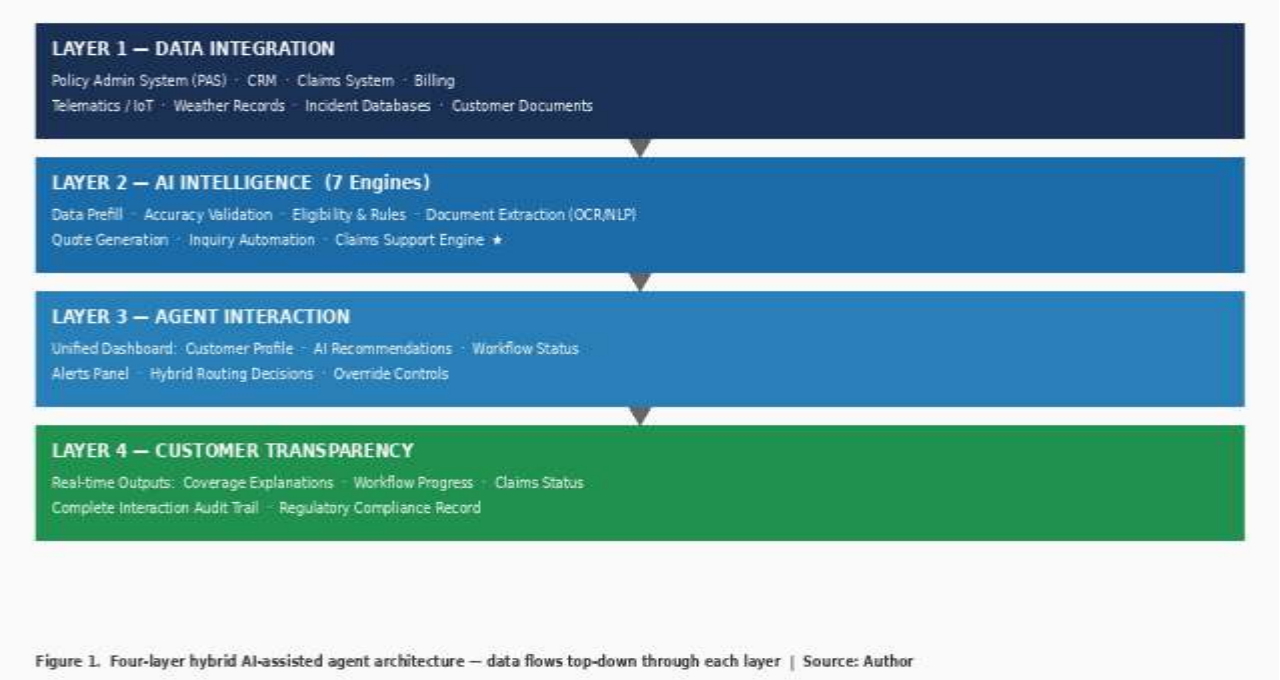

Figure 1. Four-layer hybrid AI-assisted agent architecture showing the Data Integration, AI Intelligence, Agent Interaction, and Customer Transparency layers with top-down data flow. Source: Author.

4.1. Data Integration Layer

Everything the AI Intelligence Layer works with originates here. Internal sources include the policy administration system (PAS), CRM, claims management system, billing records, and interaction history. External sources include telematics and IoT feeds, weather services keyed to the reported location and time, public incident databases covering police and emergency records, and property and vehicle reference data. Customer-submitted materials – photographs, PDFs, receipts – are ingested as they arrive during the interaction. Retrieval runs in parallel on a trigger event, bounding total latency to the slowest single source rather than the cumulative sum of all queries. Quality scores computed at ingestion travel downstream and determine whether fields can support automated processing or require agent verification.

4.2. AI Intelligence Layer

Seven engines operate here. The Data Prefill Engine populates required customer information across active workflows automatically, eliminating manual re-entry by the agent. The Accuracy Validation Engine identifies incorrect data entries before the information is entered into the record. The Eligibility and Rules Engine evaluates agent-supplied information against state underwriting regulations and the carrier's underwriting guidelines for the relevant coverage lines. The Document Extraction Engine applies OCR and NLP to raise processing accuracy for unstructured materials uploaded during the interaction. The Quote Generation Engine takes the eligibility outputs from the Eligibility and Rules Engine and incorporates them into the pricing models to create multiple quotes for the customer. The Inquiry Automation Engine handles routine follow-up communications. Section 5 covers the Claims Support Engine, the primary technical contribution. Hybrid routing governs all seven: tasks within acceptable thresholds on complexity, emotional sensitivity, and data completeness run through automated processing; others go to the agent with the partial analysis and a stated escalation reason. That routing decision is always visible to the agent, consistent with the explainability principle documented by Owens et al. (2022).

4.3. Agent Interaction and Customer Transparency Layers

The dashboard presents four panels: customer profile (policy history, coverage in force, prior claims), AI recommendations (real-time engine outputs), workflow status (task sequence and pending actions), and alerts (data quality flags, escalation notices, missing documentation). Every view is built around synthesized conclusions rather than raw data fields – the agent reads what has been determined and decides whether to act on it or override it. The Customer Transparency Layer produces the customer-facing output, transforming the result generated by AI into an easy-to-understand explanation of coverage implications and what actions are required. It also maintains an interaction record for audit trail compliance purposes, including the date of each interaction, so that the

customer can refer back to it if they have questions about any communication made to them (Owens et al., 2022; Nguyen et al., 2022).

5. Real-Time Claims Support Engine

Every experienced claims professional knows the scenario. A customer calls, clearly distressed, describing an accident or a pipe burst or hail damage. The agent has the policy record open but nothing else. The telematics data from the vehicle sits in a database the agent cannot reach mid-call. The weather report for that location at that hour requires a separate lookup. The police incident number the customer mentioned is in a public records system the agent has no live access to. So the agent takes down what they can, tells the customer they will follow up, and spends the next two hours assembling what should have been captured in the first ten minutes. The claims support engine eliminates exactly that problem.

The engine activates when a claims context is detected – through call routing metadata, agent workflow selection, or intent signals in the call audio – and runs four sequential stages. Stage 1 queries the PAS for the complete coverage picture: every active line, deductible, sublimit, exclusion, and any existing FNOL filings connected to this policyholder. That returns within seconds, before the customer has finished describing the incident. Stage 2 runs external retrieval in parallel: telematics logs for the vehicle, weather data for the reported location and time, police and emergency records matching the incident details. The agent gets independent corroboration of the account without a single follow-up call to obtain it. Stage 3 generates a question sequence using an NLP model trained on FNOL completion data and coverage documentation, calibrated to the specific coverage lines in play and the gaps left open by the external data, with each question annotated for the agent with its regulatory or coverage rationale. Stage 4 populates the FNOL package fields in real time as the agent works through the questions, delivering a complete, structured submission to the claims management system before the conversation ends. Figure 2 shows the full sequence and what it replaces.

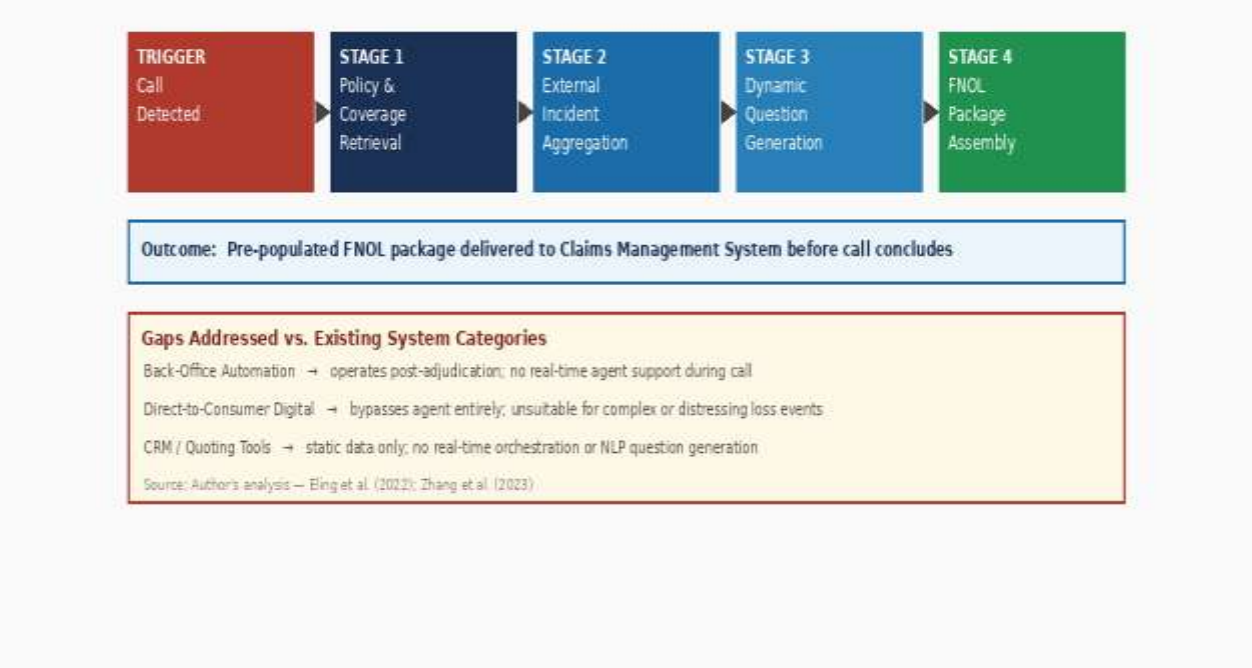

Figure 2. Real-time claims support engine: four-stage workflow from customer call trigger through FNOL package assembly, with capability gap analysis versus existing system categories. Source: Author.

6. Comparative Analysis

Table 1 measures the proposed architecture against the three existing automation categories on seven capability dimensions. The exercise is less about showing that existing systems are inadequate than about making the gap explicit: each category does what it was built for well, but none was built for the live agent-customer interaction, and combining them does not produce that capability either. The dimension set and rating procedure are described in Section 3.

Table 1. Comparison of existing insurance automation system categories vs. the proposed hybrid AI-assisted architecture

Capability Dimension	Back-Office Auto.	DTC Digital	CRM / Quoting	Proposed Architecture
Real-time agent augmentation	No	No	Partial	Yes
Multi-source data integration	Yes	Partial	No	Yes
Dynamic question generation	No	No	No	Yes – novel
Automated FNOL package assembly	Post-call	No	No	Real-time
Hybrid AI-human routing	No	No	No	Yes – explicit
Customer transparency output	No	Yes	No	Yes
Interaction audit trail	Yes	Partial	Partial	Complete

Source: Author's analysis based on Eling et al. (2022), Zhang et al. (2023), Owens et al. (2022), and Accenture (2022). DTC = Direct-to-Consumer.

The most distinctive entry in the table is dynamic question generation. Back-office fraud systems produce risk scores after submission. Digital consumer channels run decision trees without agent involvement. CRM tools can surface a prior claims history but cannot construct a question protocol matched to the intersection of specific policy terms, real-time external incident data, and state-specific regulatory documentation requirements. Assembling all three in real time during a live call is what makes Stage 3 of the engine architecturally distinct.

7. Evaluation Framework and Performance Expectations

Table 2 maps evaluation metrics across three workflow domains with baselines drawn from manual agent processes and improvement expectations grounded in published research. The 30–55% range for quoting cycle time reduction is consistent with the efficiency gains Eling et al. (2022) report for AI-augmented quoting implementations across multiple carriers. The 35–50 percentage point gain in FNOL completeness at first call reflects the documented distance between what agents capture manually during a loss report conversation and what a fully structured FNOL submission requires. Accenture (2022) identified that gap as the primary driver of post-FNOL contact cycles, which in turn drive both claims cycle time and the premium attrition risk their research quantified. All figures are literature-grounded estimates; production validation through a controlled carrier implementation is planned as the next phase of this work.

Stated formally, the framework advances five propositions. P1: deployment reduces quoting cycle time by 30–55% against a manual agent workflow. P2: automated prefill and validation raise quoting data accuracy by 15–25 percentage points over manual entry. P3: real-time eligibility checking raises underwriting submission completeness by 20–30 percentage points. P4: the claims support engine raises FNOL completeness at call end by 35–50 percentage points over the manual FNOL process. P5: complete first-call FNOL packages reduce FNOL-to-adjuster cycle time by 20–40% against standard carrier service levels. Each proposition is falsifiable in a controlled carrier implementation, and P4 – the proposition tied to the architecture's central contribution – is the primary target of the validation study described in Section 10.

Table 2. Evaluation metrics and expected performance improvements across three workflow domains

Workflow Domain	Metric	Baseline	Expected Improvement
Quoting	Cycle time reduction	Manual agent workflow	30–55%
Quoting	Data accuracy gain	Manual data entry	+15–25 percentage points
Underwriting	Submission completeness	Manual submission	+20–30 percentage points
Claims initiation	FNOL completeness at call end	Manual FNOL process	+35–50 percentage points

Workflow Domain	Metric	Baseline	Expected Improvement
Claims initiation	Cycle time: FNOL to adjuster	Standard carrier SLA	20–40% reduction

Source: Author's analysis based on Eling et al. (2022) and Accenture (2022). All improvement figures are literature-grounded estimates pending empirical validation.

8. Regulatory, Privacy, and Ethical Considerations

A system that reaches into telematics feeds, weather records, and public incident databases during a live customer call operates squarely inside the regulatory perimeter now forming around AI in insurance, and the architecture treats that perimeter as a design input rather than an afterthought. Three constraint sets matter most. First, data use and consent: telematics retrieval presupposes the consent terms of the customer's usage-based insurance program, and each external retrieval in Stage 2 is logged with its purpose, so the data trail behind any FNOL field can be reproduced on demand. The interaction record maintained by the Customer Transparency Layer serves this function for the customer; the same record serves the market-conduct examiner. Second, governance and accountability: the NAIC model bulletin requires insurers to document how AI outputs affecting consumers are generated, validated, and overseen (NAIC, 2023). The architecture's hybrid routing is deliberately conservative on this point – the AI engines synthesize and recommend, the agent decides, and every escalation carries a stated reason, which gives a governance program a per-decision audit artifact rather than a model-level attestation. Third, fairness: the question protocols generated in Stage 3 are calibrated to coverage lines and incident data, never to protected characteristics, and the training data for the question generation model requires the same bias auditing that the EU AI Act's high-risk provisions formalize for European deployments (European Union, 2024). None of this eliminates ethical risk. A system that makes claims intake dramatically more efficient could, misused, become an instrument of claim suppression rather than claim service. The design counters that by making the customer-facing output symmetrical: the same synthesis the agent sees is translated for the customer, including what the policy covers, not only what it excludes.

9. Applications and Broader Implications

Personal lines auto and property are the strongest near-term fit. The volume of first loss report calls is high, the documentation requirements are well-standardized, and the external data sources the engine depends on – telematics providers, weather services, incident databases – are sufficiently mature. For an auto claim, the engine can confirm whether telematics shows a hard braking or impact event at the reported location before the agent has asked the first structured question. For a homeowners storm claim, it can pull the National Weather Service record for that address on that date and populate the cause-of-loss field from a corroborated source rather than the customer narrative alone.

Commercial lines present a different case. The documentation burden on a workers' compensation or commercial liability claim is substantially heavier than on a personal auto loss, and the Document Extraction and Eligibility engines carry more of the efficiency argument there. For small business accounts managed by agents with mixed books, the combination of automated document ingestion and real-time eligibility checking reduces the multi-contact cycles that currently make commercial servicing labor-intensive. The deeper implication holds across both segments: when Huang and Rust (2018) found that empathetic tasks retain a durable human advantage, they were describing something carriers already observe in satisfaction data – policyholders who felt their agent genuinely listened during a claim score higher regardless of outcome. The architecture does not try to automate the listening. It removes the mechanical work that currently prevents it.

10. Limitations and Future Research

Four limitations bound the claims this paper can make. First, the architecture has not been deployed; every performance figure in Section 7 is a literature-grounded estimate, and the propositions remain untested. Second, the comparative analysis in Section 6, while structured, was conducted by a single rater from documented capabilities; an expert-panel replication would strengthen it. Third, the near-term scope is personal lines auto and property, where the external data sources are mature – commercial lines, with heavier documentation and thinner external data coverage, may realize a different balance of benefits. Fourth, the architecture's value depends on the availability and quality of third-party data feeds that carriers do not control; a telematics outage or a lagging public-records interface degrades Stage 2 to the current manual baseline. The planned next phase addresses the first two limitations directly: a controlled implementation with a mid-size carrier, instrumented to test P1 through P5; a

prototype of the Stage 3 question generation model, evaluated against FNOL documentation-completeness rubrics on historical claims data using the claims-text methods Li et al. (2025) demonstrated; and an expert panel of claims professionals to replicate the Table 1 capability ratings.

11. Conclusion

The insurance industry has spent years building automation around the agent rather than for the agent. Back-office processing has improved substantially; consumer self-service channels have expanded; and yet what happens when a policyholder calls during an active loss has barely changed. The agent is largely on their own, assembling context from systems that were designed for other purposes, while the customer waits for clarity they need immediately. The four-layer architecture described in this paper targets that specific moment. The Claims Support Engine activates during the call, retrieves and synthesizes the information the agent needs, generates a question sequence matched to the actual loss circumstances, and delivers a complete FNOL package to the claims system before the customer hangs up.

Comparative analysis across seven dimensions indicated that no existing system category – individually or combined – meets that standard. Drawing on Eling et al. (2022) and Accenture (2022), the evaluation framework projects 30–55% reductions in quoting cycle time, a 35–50 percentage-point gain in first-call FNOL documentation completeness, and a 20–40% reduction in FNOL-to-adjuster cycle time – improvements aimed directly at the claims-experience failures behind the $170 billion in premium revenue identified as at risk. The design principles are consistent with the human-AI collaboration research (Jarrahi, 2018; Huang and Rust, 2018; Duan et al., 2019), with the trust requirements documented for the insurance context (Owens et al., 2022; Nguyen et al., 2022), and with the direction of AI regulation in both the United States and the European Union. Moving from design to deployment requires the production implementation study outlined in Section 10. The Stage 3 question generation model is the most technically demanding component of that program, and recent advances in claims text analysis suggest it is within reach (Li et al., 2025).

Conflict of Interest

The author declares no conflicts of interest with respect to the research, authorship, or publication of this article.

Funding

The author received no funding for this work.

Data Availability

No datasets were generated or analyzed in this study. The comparative ratings and evaluation framework are fully reported in Tables 1 and 2.

References

1. Accenture. (2021). More consumers willing to share lifestyle data for behavior-based insurance premiums, but trust in data security falls, Accenture report finds. Accenture Newsroom. https://newsroom.accenture.com/news/2021/more-consumers-willing-to-share-lifestyle-data-for-behavior-based-insurance-premiums-but-trust-in-data-security-falls-accenture-report-finds
2. Accenture. (2022). Poor claims experiences could put up to $170B of global insurance premiums at risk by 2027. Accenture Newsroom. https://newsroom.accenture.com/news/2022/poor-claims-experiences-could-put-up-to-170b-of-global-insurance-premiums-at-risk-by-2027-according-to-new-accenture-research
3. Deloitte. (2025). 2026 global insurance outlook. Deloitte Insights. https://www.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-outlooks/insurance-industry-outlook.html
4. Duan, Y., Edwards, J. S. and Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data: Evolution, challenges and research agenda. International Journal of Information Management, 48, 63-71. https://www.sciencedirect.com/science/article/abs/pii/S0268401219300581?via%3Dihub
5. Eling, M., Nuessle, D. and Staubli, J. (2022). The impact of artificial intelligence along the insurance value chain and on the insurability of risks. The Geneva Papers on Risk and Insurance - Issues and Practice, 47(2), 205-241. https://link.springer.com/article/10.1057/s41288-020-00201-7
6. European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj
7. Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75-105. https://doi.org/10.2307/25148625

8. Huang, M. H. and Rust, R. T. (2018). Artificial intelligence in service. Journal of Service Research, 21(2), 155-172. https://doi.org/10.1177/1094670517752459
9. Huang, M. H. and Rust, R. T. (2021). Engaged to a robot? The role of AI in service. Journal of Service Research, 24(1), 30-41. https://doi.org/10.1177/1094670520902266
10. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, 61(4), 577-586. https://doi.org/10.1016/j.bushor.2018.03.007
11. Li, D., Jin, Z., Qian, L. and Yang, H. (2025). Textual analysis of insurance claims with large language models. Journal of Risk and Insurance, 92(2), 505-535. https://onlinelibrary.wiley.com/doi/10.1111/jori.70004
12. NAIC. (2023). Model bulletin: Use of artificial intelligence systems by insurers. National Association of Insurance Commissioners. https://content.naic.org/sites/default/files/inline-files/2023-12-4%20Model%20Bulletin_Adopted_0.pdf
13. Nguyen, T. M., Quach, S. and Thaichon, P. (2022). The effect of AI quality on customer experience and brand relationship. Journal of Consumer Behaviour, 21(3), 481-493. https://onlinelibrary.wiley.com/doi/10.1002/cb.1974
14. Owens, E., Sheehan, B., Mullins, M., Cunneen, M., Ressel, J. and Castignani, G. (2022). Explainable artificial intelligence (XAI) in insurance. Risks, 10(12), 230. https://www.mdpi.com/2227-9091/10/12/230
15. Peffers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45-77. https://doi.org/10.2753/MIS0742-1222240302
16. Schrijver, G., Sarmah, D. K. and El-hajj, M. (2024). Automobile insurance fraud detection using data mining: A systematic literature review. Intelligent Systems with Applications, 21, 200340. https://www.sciencedirect.com/science/article/pii/S2667305324000164?via%3Dihub
17. Zhang, W., Shi, J., Wang, X. and Wynn, H. (2023). AI-powered decision-making in facilitating insurance claim dispute resolution. Annals of Operations Research. https://doi.org/10.1007/s10479-023-05631-9