

How AI Helped Reduce Claim Review Turnaround by Days While Meeting Audit Requirements

Intertech, Inc.



How Your Software Development Team Can Make The Difference

Executive Summary

Insurance claim reviews are commonly framed as a staffing or productivity problem. In practice, it is far more often a software design problem. Most delays within claims organizations are not caused by people thinking slowly or making poor decisions. It is caused by fragmented systems, unstructured documents, missing context, and manual evidence assembly, which force reviewers to spend hours preparing to decide before they ever evaluate the claim itself.

In this case pattern, we will illustrate how software developers, partnering with claims operations and compliance teams, can redesign the internal review workflow and introduce AI at carefully chosen friction points: document understanding, deterministic validation, and evidence-linked context assembly, by building an architecture where AI accelerates preparation and comprehension while humans retain decision authority. The result is a sustained reduction in average claim review turnaround time, achieved without weakening audit controls, relaxing documentation standards, or increasing operational risk.

Turnaround Time Defined From an Engineering Perspective

Before writing a single line of new code, developers worked with operations leaders to define exactly what “turnaround time” meant in system terms. Vague cycle-time metrics are not

actionable. Engineers needed a precise flow boundary that could be instrumented.

Example of what needed to be defined:

- **Start:** Claim enters a Ready for Review state
- **End:** Claim disposition is finalized (approve, deny, or pend with documented rationale)
- **Excluded:** Time waiting on claimant responses

This definition allowed developers to separate external waiting time from internal handling time. It also enabled consistent measurement across claim types and queues.

For example, in this case, engineers added state-transition instrumentation to the claims platform. Every claim movement between states was timestamped and logged. Dashboards showed internal handling time segmented by claim category, complexity tier, and reviewer queue. Once the flow became visible, it became obvious that most of the delay occurred before substantive review even began.

The Real Bottlenecks

Workflow tracing and developer shadowing sessions revealed that reviewers were not slow at evaluating claims. They were slow to assemble the information needed to evaluate claims. The dominant time losses were technical in nature and were identified as:

- Reviewers opening multiple systems to locate documents
- Manual reading of PDFs and scanned images
- Re-keying values into structured fields
- Searching prior claim history for context
- Sending messages to request missing or unclear documents

- Redoing work when extracted values did not match system records

From a software perspective, these are classic symptoms of a missing ingestion and normalization layer, and they concluded that improving the decision interface would have a limited impact unless the data feeding that interface became reliable and structured.

Fixing Document Ingestion and Validation Pipelines

It is critical to stabilize intake. For this reason, developers built a document ingestion pipeline that sits between inbound content and the claims platform, so when documents arrive, the pipeline performs three steps:

- **First**, documents are classified by type (loss notice, invoice, medical record, police report, estimate, correspondence).
- **Second**, document AI extracts structured fields appropriate to that type.
- **Third**, deterministic validation rules compare extracted values against systems of record such as policy administration and claimant databases.

If required fields are missing, if identifiers conflict, or if scans are unreadable, the claim is flagged immediately and routed to an exception queue.

This architecture is intentionally conservative and probably not what you would expect, but to maintain control, AI is only used for extraction, while correctness is enforced by deterministic checks. In this way, uncertainty is identified early rather than being buried in downstream review. From an engineering standpoint, this converts unstructured content into a reliable, queryable data substrate that AI components can safely consume later, speeding things up without concern.

Where Developers Put AI: The Claim Brief

The largest improvement in turnaround did not come from automating decisions. It came from eliminating the time reviewers spent building context. In this way, developers implemented what became known internally as the Claim Brief.

When a claim transitions into Ready for Review, a background service retrieves data from:

- Claims system
- Policy administration system
- Document repository
- Prior claim history
- Notes and correspondence

An LLM is then used to summarize and organize this retrieved content into a structured Claim Brief that contains:

- Coverage and policy context
- Loss narrative summary
- Extracted evidence values
- Exceptions and conflicts
- Suggested next actions

The output is rendered as a structured artifact inside the claims UI. The added benefit is that reviewers start every claim with a complete, standardized baseline that, from an architecture perspective, looks like:

Systems of Record → Retrieval Layer → Prompt Template → LLM → Structured Claim Brief

Prompts, schemas, and templates are versioned, tested, and deployed through the same CI/CD pipeline as application code.

Why This Worked to Change Turnaround

Before integration, reviewers spent a large portion of their day reconstructing the story of each claim. That work was invisible in traditional metrics but consumed hours. After integration, the story was automatically assembled, allowing reviewers to move directly into evaluation and judgment rather than discovery. Multiply that time savings across thousands of claims and hundreds of reviewers, and the aggregate effect became days rather than minutes.

The software developers, using AI correctly in this case, did not make humans faster. They removed preparation labor.

Auditability Engineered into the System

As important as efficiency was the system's Audit readiness, which was treated as a first-class technical requirement rather than a compliance afterthought. Every extracted field displayed in the Claim Brief needed to include a link back to its originating document location. Every brief section had to list the documents used to produce it. Every brief generation event had to be logged with model version, prompt version, and source systems. When a reviewer edits a value or adds a note, that action was now captured. When a disposition is made, the user identity and timestamp was recorded.

This created a reproducible chain:

Source Document → Extracted Field → Claim Brief → Human Edit → Final Decision

And because of this, Auditors, to this day, can reconstruct exactly how a conclusion was reached.

Guardrails Implemented in Code

Rather than relying on policy statements, developers enforced guardrails at the service layer that included:

- AI services are not allowed to call disposition endpoints.

- Low-confidence extractions require human confirmation.
- Retrieval is restricted to approved internal systems.
- Approval actions require an authenticated human identity.

These rules were enforced through APIs, permissions, and service contracts. And to keep them safe, specific code changes were required, not casual configuration tweaks.

Encoding Review Knowledge as Runbooks

Another contributor to the delay prior to the project was reviewer variance. Different reviewers followed slightly different mental checklists. To help level the playing field, developers collaborated with claims SMEs to encode review procedures as structured runbooks, including required checks, decision trees, escalation rules, and documentation expectations per claim type. In this way, AI was used to suggest which runbook to apply based on claim characteristics, while humans executed the steps. This reduced rework caused by missed steps and increased consistency across the organization.

Why the Multi-Day Reduction Was Real

The improvement was not theoretical and not dependent on novelty. Time was removed from:

- Manual reading
- Manual searching
- Manual re-keying
- Back-and-forth clarification
- Reconstruction of claim context

None of the time savings came from skipping checks or relaxing standards. The system simply eliminated invisible labor.

In Closing

This study demonstrates a broader truth: meaningful AI impact in regulated environments comes from software engineering discipline, not model selection. The multi-day turnaround reduction happened because developers redesigned the system around evidence-first workflows and then inserted AI where it amplified human cognition. AI did not replace claims professionals. Nor did it bypass controls. Software developers simply removed friction from the system.

If you have questions or would like to continue the conversation to find out how Intertech can directly support this type of AI-enabled claims modernization across several critical areas:

- **Workflow & Turnaround Instrumentation**
Mapping current claim lifecycles, defining measurable state boundaries, and adding telemetry to expose internal handling time, queue latency, and rework sources.
- **Document Ingestion & Intelligent Extraction Architecture**
Designing and implementing pipelines for OCR, document classification, field extraction, and deterministic validation against systems of record.
- **Deterministic Validation & Rules Engines**
Building rule layers that reconcile extracted data with policy, claims, and customer systems to surface conflicts early and prevent downstream rework.
- **Retrieval Layer & Systems Integration**
Integrating claims platforms, policy administration systems, document repositories, and historical data into a unified retrieval service for AI consumption.
- **AI Claim Brief Generation**
Designing prompt templates, output schemas, and orchestration logic that produce structured, evidence-linked claim briefs rather than free-form text.
- **Human-in-the-Loop Review Experiences**
Implementing UI patterns that present summaries, citations, exceptions, and next-step recommendations while preserving human decision authority.
- **Audit Logging & Traceability**
Building end-to-end provenance tracking from source document → extracted field → brief → reviewer edit → final decision.

- **Security, Access Control & Guardrails**
Enforcing service-level constraints that prevent AI from executing dispositions, restrict retrieval to approved sources, and require authenticated approvals.
- **Runbook Encoding & Decision Support**
Translating claims SMEs' procedures into structured workflows and decision trees that AI can recommend but humans execute.
- **CI/CD for AI Artifacts**
Versioning prompts, schemas, validation rules, and model configurations as first-class artifacts alongside application code.
- **Quality Assurance & Drift Monitoring**
Designing sampling strategies, extraction accuracy monitoring, and regression tests to maintain reliability over time.

Let our team know. Intertech consultants will partner with internal IT and development teams and transform your design, build, and operationalized usage of AI in a way that measurably reduces turnaround time, strengthens audit posture, and produces systems that are maintainable long after initial deployment.

(Checklist on next page)

AI-Assisted Claims Review Checklist

*Combined Implementation
Checklist & Maturity Model*

Intertech, Inc.

To use this checklist most effectively, start by assembling a small, cross-functional group that includes software developers, an architecture lead, claims operations, and compliance or audit.

Work through each section together and mark which checklist items are fully in place today, which are partially implemented, and which do not exist. Avoid debating tooling at this stage—focus instead on observable system behavior (what the software actually enforces, logs, and produces). Once items are marked, identify the maturity column (Foundational, Operational, or Optimized) that best represents the majority of your rows; this becomes your current-state baseline.

Then, use the “Next Step” guidance to build a short, sequenced roadmap that targets the highest-impact gaps first—typically telemetry, ingestion, validation, and retrieval before advanced AI features. Revisit the checklist quarterly to measure progress, validate that improvements are producing real turnaround reductions, and adjust priorities as volume, risk profile, and regulatory expectations evolve.

Turnaround Definition & Telemetry

- Start and end states clearly defined
- External wait time separated from internal handling time
- State transitions instrumented and logged
- Dashboards by claim type and complexity

Intertech consultants work with operations and engineering teams to define precise workflow boundaries, design state models, and add instrumentation to existing platforms so turnaround becomes

a measurable, observable system property rather than an abstract KPI.

Document Ingestion & Extraction

- Centralized ingestion pipeline
- Document classification in place
- Field extraction configured per document type
- Confidence scores captured

Intertech designs and implements ingestion pipelines, integrates document AI services, and builds supporting services that normalize inbound content into structured, versioned, and monitorable data flows.

Deterministic Validation

- Extracted fields validated against systems of record
- Missing or conflicting data routed to exception queues
- Validation rules versioned and tested

Intertech builds validation layers and rules engines that reconcile extracted data with authoritative systems, ensuring correctness is enforced by software rather than manual review.

Retrieval Layer

- Unified access to claims, policy, documents, history, and notes
- Source whitelisting enforced
- Latency and error handling implemented

Intertech architects secure retrieval services that aggregate data across enterprise systems with consistent interfaces, caching, and resiliency patterns.

Claim Brief Generation

- Prompt templates versioned
- Structured output schema defined
- Evidence citations included
- Generated automatically on Ready for Review

Intertech designs prompt orchestration services, structured schemas, and background jobs that generate evidence-linked Claim Briefs as first-class system artifacts.

Human-in-the-Loop Controls

- AI cannot execute dispositions
- Low-confidence extractions require confirmation
- Human identity required for approvals

Intertech enforces human-in-the-loop patterns through API boundaries, permission models, and workflow gates rather than relying on policy alone.

Auditability & Logging

- Source-to-output traceability
- Model and prompt versions logged
- Reviewer edits logged
- Decision events logged

Intertech designs provenance models and logging strategies that make AI-assisted workflows explainable, reproducible, and audit-ready.

Runbooks & Decision Support

- Procedures documented per claim type
- Decision trees encoded
- AI suggests runbook, humans execute

Intertech helps translate SME knowledge into structured workflows and decision models that integrate directly into applications and AI services.

Quality & Drift Monitoring

- Sampling strategy defined
- Extraction accuracy tracked
- Regression tests for prompts and schemas

Intertech implements monitoring, sampling, and regression testing frameworks so AI-assisted systems remain reliable as models, data, and usage evolve.

CI/CD for AI Artifacts

- Prompts in source control
- Schemas in source control
- Automated tests
- Controlled deployment

Intertech extends existing CI/CD pipelines to treat prompts, schemas, and validation rules as versioned, testable artifacts alongside application code.

This checklist is intended as a practical diagnostic tool, not a theoretical framework. Teams that want to see meaningful turnaround improvements use it to identify concrete system gaps, prioritize engineering work that removes preparation labor, and sequence

changes so that foundations (telemetry, ingestion, validation, retrieval) are established before advanced AI capabilities are introduced. The fastest progress comes from treating this as an engineering modernization initiative with AI as an accelerator, rather than an isolated AI project.

If you would like help assessing your current state, mapping these capabilities to your existing architecture, and building a phased implementation plan, let our team at Intertech know. Our senior consultants will partner directly with your IT leaders and software development teams to design, build, and operationalize AI-enabled workflows that reduce turnaround time, strengthen auditability, and remain maintainable over the long term. Reach out to start a working session focused on your environment, your constraints, and the fastest path to measurable improvement.

Visit [intertech.com](https://www.intertech.com) for more information on our team.

Sources & Further Reading

Intelligent Document Processing

<https://adoption.microsoft.com/en-us/intelligent-document-processing/>

Multimodal Content Processing Reference Architecture

<https://learn.microsoft.com/en-us/azure/architecture/ai-ml/idea/multi-modal-content-processing>

Automated Document Validation That Auditors Trust

<https://techcommunity.microsoft.com/blog/azure-ai-foundry-blog/automated-document-validation-that-auditors-trust-the-deterministic-advantage/4394307>

AI Agents for Claims Automation: A Complete Guide

<https://www.v7labs.com/blog/automated-claims-processing-for-insurance>

Transforming the Incident Lifecycle With AI Agents

<https://www.pagerduty.com/blog/ai/transforming-the-incident-lifecycle-with-ai-agents/>

Cut MTTR by 40% Using AI for Automated Incident Triage

<https://rootly.com/sre/cut-mttr-by-40-using-ai-for-automated-incident-triage>

NIST AI Risk Management Framework

<https://www.nist.gov/itl/ai-risk-management-framework>

<https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf>

NAIC – Use of Artificial Intelligence Systems by Insurers

<https://content.naic.org/sites/default/files/cmte-h-big-data-artificial-intelligence-wg-ai-model-bulletin.pdf.pdf>

Audit logs for Copilot and AI applications

<https://learn.microsoft.com/en-us/purview/audit-copilot>

Thank you for your time.

Visit [intertech.com](https://www.intertech.com) to start the conversation.