



DOCXTRACT • WHITE PAPER 03

The Future of Claims Processing

How intelligent document processing is rewiring the healthcare claims lifecycle — from manual keying and rework toward accurate, autonomous, and near-touchless operations.

PREPARED BY
DocXtract

PUBLISHED
2026

AUDIENCE
Providers & RCM Leaders

EXECUTIVE SUMMARY

The constraint has shifted from reading documents to acting on them.

Every healthcare claim is a bundle of unstructured documents — clinical notes, prior-authorization forms, explanation-of-benefit statements, remittances, and scanned attachments — that must be read, validated, and reconciled before payment can move. For decades that work has been done by people, propped up by rules-based automation and OCR that break the moment a document strays from a fixed template.

That model is now under acute strain: denial rates are climbing, administrative costs keep rising, and staffing shortages mean work goes undone. At the same time, a new generation of intelligent document processing (IDP) — combining vision-language models, generative and agentic AI, and workflow orchestration — has crossed the accuracy threshold needed for production use. As this paper shows, the constraint in claims processing has shifted from reading documents to acting on them, and the organizations that re-architect around that shift will define the next decade of the claims lifecycle.

~\$1T

U.S. healthcare
administrative spend per year

~20%

of claims denied on average

80%

of healthcare data is
unstructured

30–60%

potential cut in cost to collect with AI

Sources: McKinsey; KFF; Healthcare Informatics Research; McKinsey RCM analysis (2026).

CONTENTS

01	The Cost of the Status Quo	04
02	Why the Old Model Is Breaking	05
03	From OCR to IDP to Agentic Document Processing	06
04	How IDP Transforms the Claims Lifecycle	07
05	The Regulatory Tailwind: CMS-0057-F	08
06	The Near-Touchless Claim	09
07	Getting Started: A Pragmatic Path	10

The Cost of the Status Quo

McKinsey estimates the U.S. spent roughly \$950 billion on healthcare administration in 2019 — 15 to 25 percent of national health spending — and concluded that about \$265 billion of it could be removed without reducing care quality, largely through simplification and automation of processes like claims. The claims-transaction layer alone runs about \$200 billion a year: the average claim costs \$12 to \$19 to process, more than 9 billion are processed annually, and a typical claim still takes four to six weeks to pay.

Denials have become a structural drag rather than an occasional exception. Analysis cited by McKinsey found nearly 20 percent of claims are denied on average, and as many as 60 percent of denials are never appealed — legitimate revenue simply written off. Experian Health's 2025 State of Claims survey found 41 percent of providers now face denial rates of 10 percent or higher, a share that has grown every year since 2022, and hospitals spent an estimated \$19.7 billion in 2022 trying to overturn denials.

Failure Point	What the Data Shows	Root Cause
Front-end data errors	~22% of denials tied to registration/eligibility errors	Manual keying, mismatched records
Documentation gaps	"Request-for-information" denials up ~70% in value (2025)	Missing or unread supporting docs
Billing accuracy	~80% of medical bills contain at least one error	Manual coding of unstructured notes
Rework economics	~65% of denied claims never resubmitted; ~\$25 per rework	Staffing limits, low ROI on manual appeal

Sources: KFF via McKinsey; MDaudit (2025); Experian Health (2025); industry billing-error analyses (2026).

The pattern is telling: most of the damage originates upstream, in how documents are read and data is captured — not in the adjudication logic. Claims processing is fundamentally a document and data problem wearing the costume of a billing problem.

Why the Old Model Is Breaking

Three pressures are converging to make the labor-plus-OCR model unsustainable.

The unstructured-data ceiling. Roughly 80 percent of healthcare data is unstructured — forms, notes, faxes, PDFs — and largely disconnected from core systems. OCR reads characters but not meaning, and on complex or handwritten content its accuracy can fall to around 60 percent. Rules-based templates hold only until a payer changes a form. The result is a permanent backlog of documents needing human interpretation — the work that doesn't scale.

The labor and cost squeeze. RCM typically costs an at-scale health system 3 to 4 percent of revenue, with health systems collectively spending more than \$140 billion a year on it. With staffing shortages, the binding constraint is now human hours, not adjudication rules. Confidence in existing tools has eroded: only 56 percent of providers believe their current claims technology is sufficient, down from 77 percent in 2022.

Regulatory acceleration. The CMS Interoperability and Prior Authorization Final Rule (CMS-0057-F), phasing in from January 2026, compels payers toward faster decisions, specific denial reasons, public metrics, and FHIR-based APIs. The compliance clock is running, and manual workflows can't meet the new timelines at scale.

"Billions have been invested in automation, yet most systems still depend on billing staff working through a patchwork maze of loosely connected point solutions."

— McKinsey & Company, January 2026

From OCR to IDP to Agentic Document Processing

For most of the last two decades, "document automation" meant OCR plus rules — scan, extract, hope the layout matched. IDP replaced fixed templates with machine learning that classifies documents, locates fields regardless of layout, and validates results. The newest generation adds large language models for context, vision-language models for layout, and agentic AI that can autonomously classify, validate, cross-reference, and route — escalating only low-confidence cases to a human.

Generation	How It Works	What Changed
OCR + rules	Character recognition on fixed templates	Fails on new layouts; ~60% accuracy on complex/handwritten docs
Classic IDP	ML classification and field extraction	Handles variation; still extraction-focused
Agentic IDP	LLM/VLM context + autonomous validate-and-route	Acts on documents, not just reads them; 99%+ best-in-class accuracy

Sources: Gartner Magic Quadrant for IDP (Sept 2025); Forrester IDP analysis (2025–26); industry benchmarks (2026).

The analyst firms have converged. Gartner published its inaugural Magic Quadrant for IDP in September 2025 — evaluating 18 vendors in a 100-plus vendor market — and its criteria have moved beyond raw accuracy toward orchestration and the ability to act on extracted data. Gartner reports 67 percent of enterprise document-processing initiatives now evaluate agentic approaches (up from 23 percent two years earlier) and forecasts 40 percent of enterprise applications will feature task-specific AI agents by the end of 2026. Forrester frames generative and agentic AI as an "equalizer" on accuracy, pushing differentiation toward workflow and governance; IDC's 2025–26 MarketScape describes the same shift toward end-to-end workflows fed by reliable document data. Claims are exactly the high-volume, rules-governed, document-heavy workflow this technology is built to absorb.

How IDP Transforms the Claims Lifecycle

IDP doesn't just speed up existing steps; it changes where human effort is spent. Machines take on the high-volume, rules-bound reading and validation, while people move to exceptions, judgment, and patient-facing work.

- **Intake and classification.** IDP ingests documents of any format — handwritten, scanned, or digital — and classifies and normalizes them before anything reaches a human queue, attacking the front-end errors behind roughly a fifth of denials.
- **Extraction and validation.** Best-in-class IDP reaches 99 percent or higher extraction accuracy versus ~60 percent for OCR on difficult content, and attaches a per-field confidence score and audit trail to every extraction — the traceability that makes autonomous processing safe in a regulated setting.
- **Denial prevention.** Clean, validated data lets claims be checked against payer rules before submission. Among the 14 percent of providers already using AI in claims, 69 percent reported it reduced denials or improved resubmission success; 67 percent of providers believe AI can improve the process — belief now well ahead of adoption.

On the back end, McKinsey's 2026 analysis argues that AR follow-up, underpayment and denials management, and cash posting are the ideal starting point: high-volume, rules-governed, and lower-risk than patient-facing steps. It estimates AI enablement of the revenue cycle could cut cost to collect by 30 to 60 percent — for a health system with \$6 billion in patient revenue, shaving one to two points off a 3.5–4.0 percent cost-to-collect could mean \$60 million to \$120 million in savings.

99%+

best-in-class IDP extraction
accuracy

69%

of AI users report fewer
denials

35–50%

cost cut in doc-intensive
back office*

*Gartner-referenced hyperautomation outcomes. Sources: IDP benchmarks (2026); Experian Health (2025).

The Regulatory Tailwind: CMS-0057-F

What makes this moment different is that regulation now pulls in the same direction as the economics. The CMS Interoperability and Prior Authorization Final Rule (CMS-0057-F), finalized in 2024 and phasing in from January 2026, applies to Medicare Advantage, Medicaid and CHIP programs and managed care entities, and Qualified Health Plan issuers on the federal exchanges.

Requirement	What It Means	Effective
Faster decisions	72 hours (expedited) / 7 calendar days (standard)	Jan 1, 2026
Denial transparency	Payers must give a specific reason for each denial	Jan 1, 2026
Public metrics	Prior-authorization data reported publicly (CY2025 first)	By Mar 31, 2026
FHIR-based APIs	Electronic prior-authorization and data-exchange APIs	Jan 1, 2027

Source: CMS, Interoperability and Prior Authorization Final Rule (CMS-0057-F).

THE PRACTICAL CONSEQUENCE

CMS projects the rule could save roughly \$15 billion over ten years. Manual, fax-and-phone workflows cannot meet a 72-hour expedited window at volume, which requires ingesting, reading, and routing supporting documentation in near real time. The rule turns document automation from a cost-saving option into an operational necessity.

The Near-Touchless Claim

McKinsey describes the destination as a "touchless" revenue cycle — an interconnected network of agents from intake to collections, with a human always in the loop to train models, work exceptions, and ensure compliance. The transition is incremental, unlocking first in discrete back-end steps, then connecting end to end.

Horizon	Dominant Mode	What Changes for Claims
Today	Assisted	IDP removes manual keying; humans drive adjudication and appeals
Near term (1–2 yrs)	Semi-autonomous	Agents work high-volume back-end queues; staff handle exceptions
Emerging (2–3 yrs)	Near-touchless	Connected agents span intake→collections; humans supervise and govern

Synthesis of McKinsey (2026), Gartner and Forrester agentic-AI research. Timelines are directional.

For payers, the largest opportunity is in claims automation, fraud detection, and care-management analytics — McKinsey estimates private payers could save \$80–110 billion annually through AI. For providers, gains show up as lower cost to collect, faster cash, fewer preventable denials, and staff redirected to higher-judgment work. For patients, the effect is faster access and simpler billing.

Two cautions temper the optimism. Forrester notes that ROI and governance concerns will keep most organizations on deterministic automation through 2026 even amid vendor pressure — the shift is real but uneven. And McKinsey warns against "pilot purgatory": proofs of concept need pragmatic success metrics and fast scaling, with as much investment in change management and reskilling as in the technology. The winning pattern is human-plus-agent, not human-versus-agent.

Getting Started: A Pragmatic Path

Five recommendations for payer and provider RCM leaders evaluating where to start:

- 1 Start where risk is low and volume is high.** The administrative back end — AR follow-up, denials and underpayment management, cash posting — is the natural first deployment.
- 2 Sequence use cases; don't boil the ocean.** Automate one high-value workflow, prove value, then build on it. Concentrating effort within RCM creates value faster than scattered pilots.
- 3 Insist on accuracy, traceability, and governance.** Look for per-field confidence scoring, audit logs, and bounded autonomy that escalates low-confidence cases — the capabilities Gartner and Forrester now treat as differentiators.
- 4 Prioritize healthcare-specific understanding.** The system must handle handwritten, scanned, and multi-format files, interpret clinical terms and payer policies, and integrate with core adjudication and RCM systems.
- 5 Invest in people.** Plan for reskilling so staff move into exception handling, model training, and higher-judgment work.

WHERE DOCXTRACT FITS

DocXtract is an AI-powered document processing platform built for exactly the challenge this paper describes: turning the unstructured documents behind every claim — clinical notes, prior-authorization forms, EOBs, remittances, and scanned attachments — into clean, validated, structured data that flows straight into adjudication and RCM systems. With healthcare-aware extraction, per-field confidence scoring, human-in-the-loop exception handling, and integration into core systems, it helps payers and providers cut manual keying, prevent avoidable denials, and meet tightening regulatory timelines — the practical first steps toward a near-touchless claims process.

Book a demo →

docxtract.io | connect@docxtract.io

SOURCES & REFERENCES

1. McKinsey & Company. "Agentic AI and the race to a touchless revenue cycle." January 2026.
2. McKinsey & Company / Nikhil Sahni et al. "Administrative simplification: How to save a quarter-trillion dollars in US healthcare." 2021; and related JAMA Viewpoint.
3. "Active steps to reduce administrative spending associated with financial transactions in US health care." PubMed / McKinsey Center for US Healthcare Improvement, 2023.
4. KFF (Kaiser Family Foundation). "Claims denials and appeals in ACA Marketplace plans in 2023." January 2025.
5. Experian Health. "2025 State of Claims" survey (250 healthcare professionals). September 2025.
6. MDaudit. Annual denials and audit report. November 2025.
7. Gartner. "Magic Quadrant for Intelligent Document Processing Solutions." September 2025; "From Data to Decisions: Future-Proofing IDP with AI Agents," April 2026.
8. Forrester Research. "The Forrester Wave: Document Mining and Analytics Platforms, Q2 2026"; "Predictions 2026: Automation at the Crossroads."
9. IDC. "MarketScape: Worldwide Intelligent Document Processing Software 2025–26 Vendor Assessment."
10. Centers for Medicare & Medicaid Services (CMS). "Interoperability and Prior Authorization Final Rule (CMS-0057-F)." 2024–2026.
11. Healthcare Informatics Research / Applied Clinical Trials, on unstructured healthcare data.
12. Industry analyses of medical-billing and data-entry error rates (2025–2026).

© 2026 DocXtract. For informational purposes only; reflects third-party research current as of publication. Analyst forecasts and market figures are estimates and may vary across sources and over time. Not legal, financial, or compliance advice. Gartner, Forrester, McKinsey, IDC, and CMS do not endorse any vendor, product, or service.

DocXtract