SOLUTION BRIEF · HEALTHCARE & TPA

Every claim document, read and adjudication-ready.

DocXtract converts discharge summaries, hospital bills, prescriptions and diagnostic reports into validated structured data — so TPAs and insurers settle faster with fewer denials and a cleaner audit trail.

Claims remain the most document-intensive process in health insurance, and the paperwork is the bottleneck. **Around 80% of healthcare data is unstructured —** clinical notes, scanned bills, handwritten prescriptions **—** and industry analysis attributes roughly **42% of claim denials to incorrect or incomplete data entry**. Each denied claim costs **₹2,200 to ₹16,000 to rework**, and **65% of denials are never resubmitted at all**.

The processing overhead compounds it: **manual claim handling costs about ₹550 per claim versus ₹185 automated**, and reviewers spend **15 to 25 minutes on a single claim**. Automating the document layer is now the standard first move **— 68% of insurers reported reduced turnaround with AI-based claims processing**, and **86% now use AI in claims workflows**.

BUILT FOR

- → TPA claims and pre-auth desks
- → Health insurer operations and medical audit
- → Hospital billing and revenue cycle teams
- → Fraud, waste and abuse analytics

42%	₹550→₹185	68%	98%+
Of claim denials traced to data-entry errors	Cost per claim, manual vs. automated processing	Of insurers report faster turnaround using AI in claims	Field-level extraction accuracy with DocXtract

Where the claim gets stuck

01 Every hospital submits differently

A network of thousands of providers means thousands of bill formats, itemisation conventions and abbreviations. Template-based capture requires a configuration per hospital; DocXtract reads layout and meaning together, so a new provider needs no setup.

02 Clinical detail sits in prose

Diagnosis, procedure, length of stay and comorbidities are written into discharge summaries as narrative, often handwritten. Adjudication rules cannot run until that narrative becomes fields — the step that consumes most of a reviewer's 15–25 minutes.

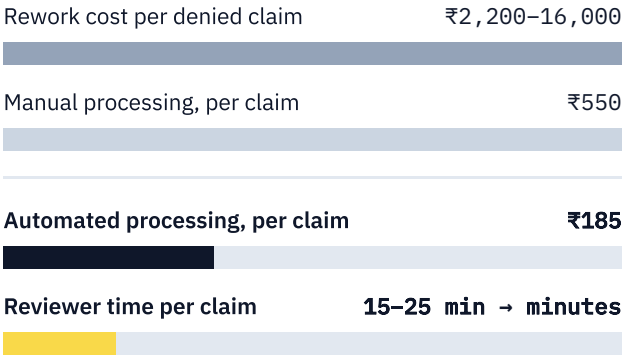
03 Denials caused by the intake, not the claim

Mismatched policy numbers, missing pre-auth references and unreadable line items produce avoidable rejections. With 42% of denials rooted in data entry and 65% never resubmitted, the loss is realised at intake, not adjudication.

04 Fraud signals only visible across documents

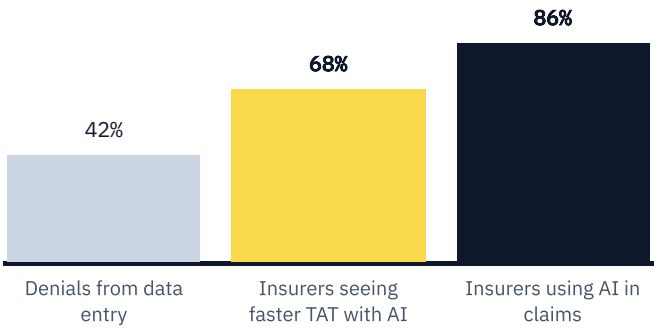
Duplicate billing, upcoded procedures and dates that contradict the discharge summary surface only when bill, summary and pharmacy list are compared as structured data. Manual review rarely reaches that comparison at volume.

FIG. 1 — COST OF A CLAIM, AND OF GETTING IT WRONG



Per-claim costs and reviewer time per Docsumo and Experian Health analyses, as reported 2026; USD converted at ₹88/$1. Bar lengths are indicative.

FIG. 2 — WHAT AUTOMATION CHANGES



Denial cause per Experian Health / industry analysis; AI adoption and turnaround figures per Docsumo and Sagacity insurance-automation research, 2026.

Document coverage and validation

STAGE	DOCUMENTS EXTRACTED	BUILT-IN VALIDATION
Pre-authorisation	Pre-auth request forms, treating-doctor notes, estimate letters, policy and ID copies	Policy-number and member match, coverage and sub-limit checks, missing-document flags before issue
Clinical record	Discharge summaries, admission notes, operation theatre and ICU records, handwritten prescriptions	Diagnosis and procedure capture, admission/discharge date consistency, length-of-stay reasonableness
Billing	Itemised hospital bills, pharmacy invoices, implant invoices, payment receipts	Line-item totalling, tariff and package comparison, duplicate-line and duplicate-claim detection
Diagnostics	Lab reports, radiology and pathology reports, investigation summaries	Test-to-diagnosis relevance, date sequencing against admission, unbilled-investigation flags
Settlement & audit	Claim forms, bank and KYC documents, deduction and denial letters	Field-level confidence scoring; low-confidence and clinically ambiguous fields routed to a medical reviewer, never guessed

How it deploys

1. **Ingest where claims already arrive.** Hospital portal uploads, email, courier scans, provider API — no change to how hospitals or members submit.
2. **Extract without templates.** Model selected per document type; JSON returned in real time at up to 5,000 pages per hour, including handwritten and mixed-language content.
3. **Validate before adjudication.** Policy matching, cross-document consistency and duplicate detection run inside the platform, so only clean claims reach the rules engine.
4. **Post to the claims system.** Pre-built connectors for SAP, Oracle, Tally, Zoho, UiPath and Power Automate, plus REST API into the TPA or insurer core.

Controls medical audit asks for

Every extracted field carries a confidence score, a reference back to the source page and a timestamped record of who reviewed it — so a deduction or denial can be defended against the document it came from, months later.

- Field-level audit log and version history
- Configurable human-in-the-loop thresholds
- Role-based access; data residency options
- Live in days; 1-day integration

NEXT STEP

Run a live claim file through DocXtract.

Send a de-identified sample set — discharge summary, itemised bill, pharmacy invoices and diagnostics. We report field-level accuracy and projected touchless rate against your current settlement TAT.

Book a demo — docxtract.io

Request a free trial · No commitment

SOURCES

1. Docsumo, "Insurance Claims Processing Automation", 2026 — manual $6.20 vs. automated $2.10 per claim (≈₹550 vs. ₹185 at ₹88/$1); 15–25 minutes reviewer time; 68% of insurers reporting reduced turnaround; 86% AI use in claims workflows.
2. Experian Health / industry denial analysis, as reported 2026 — approximately 42% of claim denials attributable to incorrect or incomplete data entry.
3. Healthcare revenue-cycle research, as reported 2026 — $25–$181 (≈₹2,200–16,000) rework cost per denied claim; roughly 65% of denied claims never resubmitted.
4. Industry estimates of unstructured healthcare data, widely cited at approximately 80% of clinical and administrative records.
5. Sagacity and insurance-automation research on AI adoption in claims operations, 2026.
6. DocXtract product documentation and benchmarking, docxtract.io, 2026 — accuracy, throughput, connectors, deployment timelines.

Figures are industry benchmarks, not DocXtract client results, except where DocXtract is named. Verify before use in regulated collateral.