ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Building Intelligent Claims and Data Infrastructure with AI

Christopher Wilson
Asst Professor

Abstract

An abstract encapsulates the essence of a research work. This piece presents an overview of AI-Enabled Payer Intelligence in healthcare. Schmitt and Gamasai explore the processing of payer claims—an activity that defines each health insurance company in the United States—through an automated claims and data pipeline. A data integration strategy, stakeholder data-sourcing choice, supportive enterprise data governance, and the automation of end-to-end operations in a controlled manner allow for greater efficiency and lower costs. Automated management of fraud, waste, and abuse; natural language understanding of adjudication; and predictive modeling applied upstream, upstream, midpoint, and downstream in supplementing human judgment throughout the operations chain are pursued.

Automated processing of claims constitutes an essential component of the operation of any North American healthcare payer. Claims-management autoroutes; a natural-language-understanding model dedicated to adjudication; and prompt-creation, natural-language-generation, and predictive-modeling approaches applied throughout the claims lifecycle down a human-in-the-loop pathway support greater efficiency, lower costs, enhanced quality, and improved experiences.

Keywords: AI-Enabled Payer Intelligence, Healthcare Claims Processing, Automated Claims Management, Claims Adjudication Systems, Fraud Waste and Abuse Detection, Healthcare Data Pipelines, Enterprise Data Governance, Predictive Claims Analytics, NLP in Healthcare, Natural Language Understanding, Claims Lifecycle Automation, Human-in-the-Loop AI, Healthcare Cost Optimization, Intelligent Claims Routing, Data Integration in Healthcare, AI in Health Insurance, Operational Efficiency in Payers, Healthcare Analytics Systems, Decision Support in Claims, AI-Driven Healthcare Operations.

1. Introduction

Healthcare costs in the United States, estimated at 3.8 trillion dollars in 2020, are primarily funded by three parties: the government (60%), private employers (31%), and individuals (9%). Consequently, out of every dollar spent on healthcare services, 89 cents are

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

channeled to providers and pharmaceutical manufacturers, whilst the remaining 11 cents are retained by insurance companies—costing them nearly 42 billion dollars annually. The process of converting health insurance premiums into reimbursements for healthcare services is largely manual, thereby creating a highly intricate and expensive procedure that is fraught with inefficiencies, fraud, and abuse.

Claims processing expenditures can reach at least 10% of the claims' payment value. Although artificial intelligence (AI) has been implemented in certain aspects of claims processing, the desire for lower costs and increased data-driven insights has ushered in the age of Payer Intelligence. Data pipelines regulated by AI models govern automation across all claims processing-related functions whilst feeding advanced analytics, natural language processing services, predictive models, and machine learning solutions. Such implementations focus on small subsets of the claims process, whereas the objective is to advance beyond piecewise automation into a fully end-to-end-enabled Payer Intelligence model capable of combining claims automation with holistic business intelligence, predictive modeling, and open-domain natural language understanding.

Fig 1: Payer Ecosystem

2. Foundations of Payer Intelligence

Defining payer intelligence is not easy. The term Payer Intelligence has been loosely used in publications, spoken about at conferences, and in the marketing outreach of artificial

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

intelligence (AI) and analytics vendors. A foundational tenet of AI and machine learning (ML) is that the approach must be data-driven: data is the oil of the new economy. Data helps build consuming personas to deliver curated advertising to pull consumers down the funnel. In healthcare, the focus has been on demand-side intelligence—predicting in hospital admissions, emergency department use, and prescription behavior—that leads to marketing, outreach, and improving care management for better member experience and quality of care as well as lowering costs. All of these initiatives fall short in the predicted value because much of the value in the healthcare dollar is being made for health plans in the claim processing cycle of revenue collection. Demand is understood. Supply and operations have not been optimized.

For healthcare payers, the supply-side intelligence required to create a demand-side capability is for a data pipeline that reduces costs, increases efficiency, and ensures quality of care. Claims have traditionally been considered standardized products to be processed with inexpensive labor. Initiatives to “automate” claims processing through hand-off processes have existed for decades, but scale will only come when automation does not require human intervention in the majority of the claims within the payer operational framework, not that of a vendor or partner. Most process workflows remain unchanged, requiring processing and adjudication decisions to be made except for a few rule-based claims. The economics of the current claims processing approach are clear: the cost of processing a claim continues to exceed the cost of the claim itself.

2.1. Definitions and scope

Payer intelligence is defined as the ability to derive actionable insights from the complex and often unwieldy data generated by health insurance claims. Claims processing—admiration and rejection decisions, data exchanges, operational performance management, predictive modeling, estimation, pricing, and underwriting—and related processes (in particular, population health analytics, provider cost prediction, and fraud, waste, and abuse detection) are often considered prudent initial targets for automation. The processing of health insurance claims is indeed a data-rich domain characterized by a multitude of repetitive, rules-based, and, in many cases, bespoke processes. Automated handling at scale stands to deliver significant financial savings through cost reduction and increased operational effectiveness.

The implementation of automated claims processing and related solutions by health insurance companies constitutes a clear application of payer intelligence. Indeed, the operational risks associated with the COVID-19 pandemic added impetus to existing efforts for automating claims processing and supporting operations in general, although the response of the health

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

insurance industry to the COVID-19 pandemic remains surprisingly immature given the economic and social magnitude of the crisis. Initiatives in health insurance claims processing automation have primarily been spearheaded by technology vendors offering point solutions, often relying on non-native automation tools without calling out—and therefore, underselling—the strong payer intelligence capability of their solution offerings. In most cases, they do not consider the end-to-end configuration of an operation, often focussing on isolated problem sections addressed through localized solutions rather than continual automation of the entire claims life cycle.

Table 1. Core numerical statements

Quantity	Value from article	Meaning
Total U.S. healthcare cost (2020)	$3.8 trillion	Total healthcare spending base
Government funding share	60%	Public contribution
Private employers funding share	31%	Employer contribution
Individuals funding share	9%	Direct household contribution
Share going to providers + pharma	89%	Payment side of the healthcare dollar
Share retained by insurers	11%	Administrative / retained side
Annual insurer-retained amount	~$42 billion	Article's stated retained amount
Claims processing expenditure	At least 10% of claims payment value	Administrative burden benchmark
AI automation cost reduction	15%–30%	Payer case-study range
AI automation throughput improvement	40%–60%	Article's cross-industry benchmark

2.2. Historical evolution of claims processing

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Claims processing has evolved from a paper-based activity to the adoption of enterprise resource planning (ERP) solutions that automate back-office functions for health insurance claims processing. While early systems focused primarily on operational efficiency and cost reduction, recent developments in AI enable financial and business analytics capabilities over an insurer's claims and other back-office data. Important operations-related decisions are therefore being automated for specific initiatives, such as reducing fraud, waste, or abuse in claims processing, improving subrogation or recovery analysis, or optimizing contract negotiation with medical providers.

Payer intelligence broadens foresight and decision-making to inform strategic and tactical operational choices via predictive models—offering a step beyond ERP systems by allowing insurers to look inward for insights rather than relying solely on outward-facing technologies, such as claims investigation and anti-fraud monitoring systems employed by government agencies. Many players are developing these predictive models to deploy in the specific context of individual initiatives, but, on their own, they represent only pockets of analytic capability without a strategic overlay. The next step is therefore to establish integrated enterprise data pipelines that will provide the base for payer-led automated data and analytic pipelines that comprehensively span all lines of business, internal functions, and planning horizons, enabling deeper and broader foresight over costs and capabilities.

Equation 1. Total funding decomposition

The total healthcare spending is funded by:

- Government = 60%
- Private employers = 31%
- Individuals = 9%

Let total spending be

$$H = 3.8 \text{ trillion dollars}$$

Then the funding shares must sum to the total:

$$H = H_g + H_e + H_i$$

where

$$H_g = 0.60H, \quad H_e = 0.31H, \quad H_i = 0.09H$$

Substitute $H = 3.8$:

$H_g = 0.60 \times 3.8 = 2.28$ $H_e = 0.31 \times 3.8 = 1.178$ $H_i = 0.09 \times 3.8 = 0.342$

Therefore,

$$H = 2.28 + 1.178 + 0.342 = 3.8 \text{ trillion}$$

Final form

$\boxed{H = 0.60H + 0.31H + 0.09H}$ $\boxed{3.8 = 2.28 + 1.178 + 0.342}$

3. Methodology

Two complementary lenses are employed to explore ways healthcare payers can leverage functional automation and gain operational efficiencies. The first focuses on data pipelines, offering insights into challenges and leading practices for implementing pipelines that fulfill a healthcare payer’s data requirements. The second lens zeroes in on claims automation, delineating how natural language processing (NLP) and machine learning (ML)—coupled with standard automation tools—can support the full claims life cycle from intake to remittance. Doing so would significantly reduce costs and processing times relative to payer-supported initiatives, while also boosting data accuracy and quality across key services.

Payer intelligence applications encompass a rich span of advanced-analytics, machine-learning, and automation solutions that offer both operational-enablement and decision-support capabilities. The spectrum of potential use cases is enhanced by the integration of emerging technologies such as natural-language understanding for frontend or backend claims adjudication; speech-to-text for transcripts of telephone conversations; predictive modeling support for FWA investigations; or near real-time detection of outlier patients and providers for focused patient outreach, network steering, and provider-reimbursement negotiations. Naturally, the broader opportunity set covered by the definition in Figure 1 is not meant to be exhaustive, nor is it intended to suggest that all advertiser-support roles need be addressed.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

3.1. Data Acquisition and Integration Strategies

Incorporating artificial intelligence into data pipelines enables them to address varied business needs for modern payers. Potential limitations include high data volumes, challenging integration and quality issues, changing regulatory requirements, and increasing concerns regarding data privacy. Crucial aspects of data pipelines encompass (i) a flexible data architecture that accommodates diverse sources and targets; (ii) strategies for effective data integration, quality management, and governance; and (iii) precise data documentation, discovery, and policy-enforcement mechanisms. Implementing such components will prepare payers to leverage pipeline capabilities and promote data-driven AI.

Data pipelines acquire and prepare data from diverse internal and external sources, making it available to stakeholders without needing data engineering assistance. By assuring data cleanliness, accessibility, and security, these pipelines allow organizations to become data-first enterprises. For healthcare payers, the primary internal sources include policies, claims, membership, and payments, all of which provide sensitive information. Similarly, third-party suppliers, government agencies, and health-maintenance organizations can supply business-critical data. Importantly, the adoption of cloud technologies relieves payers of the burden of managing their data infrastructure.

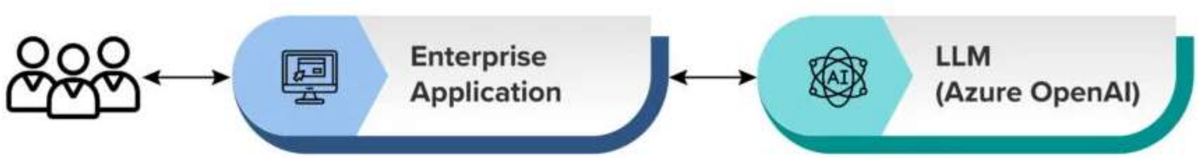


Fig 2: AI is Transforming Payer Operations

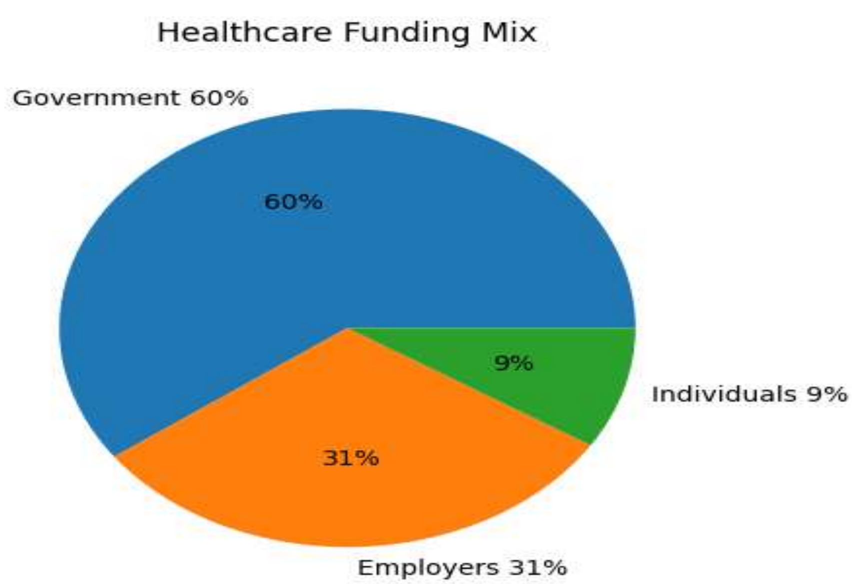
4. Objective of the Study

Artificial intelligence (AI) technologies are gradually penetrating multiple healthcare payers' operations, beginning with areas of high transaction volumes and inefficiencies. Claims processing appears as one of the first such candidate areas: demand for end-to-end automated processing has surged in recent years as claims backlogs lengthened, resource shortages grew, the costs of processing challenged profitability, and new entrants with automated capabilities gained market share. Indeed, across industry, claims automation has increasingly become a

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 01 ISSUE: 01
RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

prerequisite for continued market participation. Yet, progress has remained limited and piecemeal, concentrated in specific processes rather than encompassing the entire claims lifecycle.

The current contribution, resulting from a multiyear research effort, reconciles this expressed market desire for end-to-end claims lifecycle automation with the prevailing limited implementations. It introduces a novel data-pipelines framework, encompassing data acquisition, quality, governance, and integration, as a prerequisite foundation to underpin successful automation for the claims segment of healthcare payer operations. More specifically, concrete data-pipeline recommendations are presented to facilitate AI-enabled payer intelligence.



4.1. Research Goals and Anticipated Outcomes

The primary objective is to establish a systematic, objective understanding of AI-Payer Intelligence and the associated operational impacts. AI-Payer Intelligence refers to the combination of intelligent technology platforms and operating models deployed by payers for

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

automating operations, with a specific focus on claims management. An overview and categorization of these technology-enabled data pipelines and automation solutions—along with the underlying principles, enablers, and constraints—provide a holistic perspective of AI-Payer Intelligence.

This foundational understanding caters to a diverse audience: policy makers assessing the economic implications of AI-based automation; payers contemplating the automation of operations, such as claims management; healthcare providers and institutions concerned about the quality and costs of outsourced operations; automation technology vendors eyeing expansion in payers; and AI/ML researchers uncovering new applications of AI. Both predictive models for fraud, waste, and abuse and natural understanding solutions to read member communication have attained maturity within leading payer organizations. Granting the health plan direct operational exposure to such state-of-the-art solutions enables automation of the end-to-end claims life cycle.

5. Research Summary

Healthcare payers collect and process vast amounts of data to support critical operations. Data pipelines streamline data acquisition and integration, enhance data quality and standardization, and establish governance to ensure proper use. New generative and predictive AI capabilities are revolutionizing claims processing by combining natural language understanding to adjudicate claims automatically and predictive models for detecting fraud, waste, and abuse. With AI performing much of the work, workflow automation enables end-to-end digitization and release of resources from low-value tasks.

Payer operations are thus undergoing a major transformation. The renewed focus on automation and operations transformation promises to deliver significant benefits—cost reductions by improving operational efficiency, enhancing quality of care and member experience, and decreasing charges and premiums—yet comes with some risks owing to privacy concerns of using data and the consequences of failures. Failure to address such concerns can result in reputational and financial damage, including voluminous claims and substantial fines due to regulatory breaches. Businesses must therefore consider data governance and compliance requirements (including the consequences of poor-quality data) when defining automation strategies. Several implementation readiness assessments and roadmaps have been proposed to guide these strategic decisions.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Equation 2. Healthcare-dollar allocation equation

The states that, for every healthcare dollar:

- 89 cents go to providers and pharmaceutical manufacturers
- 11 cents are retained by insurance companies

Let one healthcare dollar be

$$D = 1$$

Let

- P = provider/pharma share
- I = insurer-retained share

Then

$P = 0.89D \ I = 0.11D$

Since all of the dollar is allocated,

$$D = P + I$$

Substitute:

$1 = 0.89 + 0.11 \ 1 = 1$

Final form

$$\boxed{D = 0.89D + 0.11D}$$

For $D = 1$,

$$\boxed{1 = 0.89 + 0.11}$$

5.1. Overview of Data Pipelines for Healthcare Payers

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Data pipelines are crucial for automating administrative operations across the healthcare sector, converting essential but cumbersome business functions into scalable operations that generate timely and trusted data and insights. They ingest raw data from a multitude of internal and external sources, including structured claims data and information locked within clinical and social services data, and apply processes that cleanse, standardise, enrich, and expose them to operational consumers. These automated data pipelines can dramatically improve services across healthcare payers, hospitals, and care providers. Within healthcare payers, for example, they enable ad hoc analytics, support regulatory compliance needs such as risk adjustment and value-based purchasing programmes, cut costs and enable faster decision making on fraud, waste, and abuse by targeting known issues, and provide the bedrock for pioneering NLP and machine learning solutions.

While the introduction of AI to enable claims automation is of great interest to many, it will only bring value when all elements of the claims lifecycle—from intake through adjudication, review, and outsource—are transformed using a human-centred design approach. As with any AI application, the economic impact will differ based on circumstance and stakeholder actions. Non-payer-led automation implementations are also being proposed. Regardless, all players in the ecosystem should be aware of the technology, its potential, and its challenges to support their own plans—those preparing for adoption should engage in external benchmarking. For a thorough consideration, attention is drawn to all aspects of the technol-ogy transformation, including data preparation and quality control, the setup of tools for machine learning model creation and deployment, the challenge of legally addressing bias in data and models, the establishment of human-in-the-loop setups for exception handling, the integration of patient consent and privacy requirements, the longer-term need for regulatory-driven auditability of decisions, and the essential attention to data governance after solution deployment.

Table 2. Derived values from the article's numbers

Derived quantity	Formula	Result
Government-funded amount	0.60 × 3.8T	$2.28T
Employer-funded amount	0.31 × 3.8T	$1.178T
Individual-funded amount	0.09 × 3.8T	$0.342T
Providers/pharma amount	0.89 × 3.8T	$3.382T
Insurer-retained amount from 11% rule	0.11 × 3.8T	$0.418T = $418B

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Derived quantity	Formula	Result
Processing cost ceiling proxy	0.10 × 3.382T	$0.3382T = $338.2B
Processing cost per $1 spend	0.10 × 0.89	$0.089 = 8.9¢
Remaining insurer-retained amount after processing	11¢ − 8.9¢	2.1¢ per $1

6. Data Pipelines in Healthcare Payers

Data pipelines serve as the central nervous system for healthcare payers, enabling the ingestion, storage, retrieval, and transformation of data in support of operational strategies, business objectives, and regulatory requirements. These pipelines generally encompass three key areas of focus: data sources and integration, data quality and standardization, and data governance and oversight. By supporting high-quality, high-fidelity data—which are fundamental requirements for success—consolidated data pipelines can be readily leveraged for a wide variety of downstream use cases that drive value across the organization.

Healthcare payers are a particularly data-rich segment of the healthcare industry. As health insurers, they control claims processed for medical services, which are consolidated into member- and provider-facing documents; financial transactions for care provided, which are integrated into provider-billing systems; and premium revenue received from customers, which fund these other activities. As large, multidimensional datasets are operationalized and made available internally for use cases such as business intelligence, claim analytics, and actuarial services, companies by design also gain collateral benefits from greater data management capabilities in privacy-preserving lossless deidentification for external sharing and compliance-oriented audit trails and controls, such as those mandated under the Health Insurance Portability and Accountability Act of 1996.

6.1. Data sources and integration

The automated claims and data pipelines leverage both internal and external data sources. Internal sources primarily consist of service utilization and patient outcomes data from healthcare providers, while external data incorporates socioeconomic, lifestyle, and demographic information, along with provider and health plan network datasets. Automated data extraction from these diverse sources serves to satisfy the data inquiry requirements of artificial intelligence (AI)-enabled claims automation and other operational workflows across multiple business

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 01 ISSUE: 01
RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

functions within a payer organization. Comprehensive data preparation creates a data foundation for AI workloads, advanced analytics, and supporting operational functions.

Data integration across business silos addresses the data availability and accessibility challenge for data analytics, data science, business intelligence, and AI workloads across the payer ecosystem. Data is integrated into a single enterprise data repository from various transactional and operational systems to deliver a consistent, complete, high-quality, and trustworthy view of the internal business operations of the payer organization. Data pipelines enable comprehensive and efficient enterprise data governance, control, and stewardship in supporting a broad range of regulatory and compliance activities, including but not limited to, Form 990, CCSI, HEDIS, NCQA, and SOX regulations.

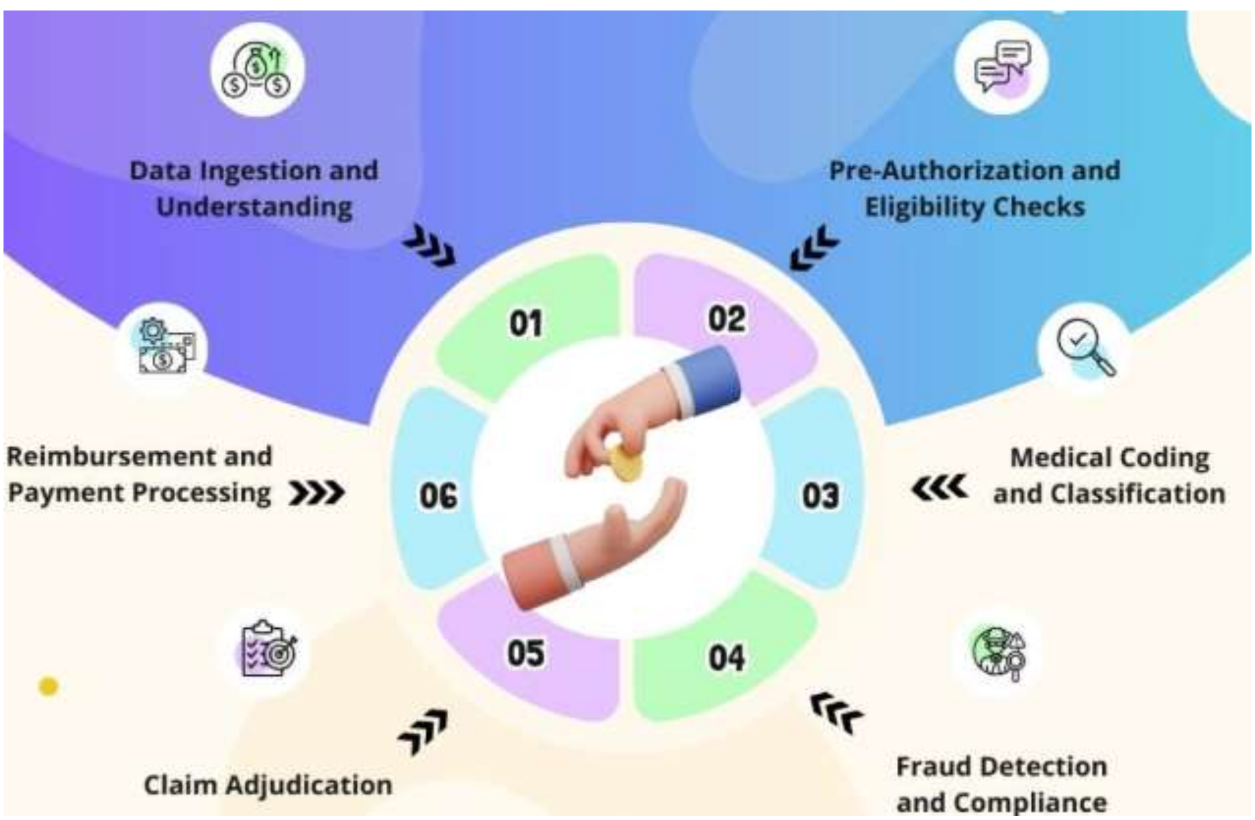


Fig 3: Agentic Process Automation in Healthcare Claims

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

6.2. Data quality, standardization, and governance

Data integration is often an art more than a science, and the challenge is magnified in most enterprises by data quality issues. Dirty data—data that is incomplete, inconsistent, inaccurate, or excessively noisy—costs organizations trillions annually. Applications powered by dirty data suffer from decision fatigue and a lack of actionable insights, data scientists spend inordinate amounts of time cleaning up data, campaigns based on dirty data fail, and quality control becomes more of an afterthought than a deliberate, early, ongoing step. Indeed, dirty data can even put organizations at risk of regulatory violations.

A holistic and comprehensive data quality and governance strategy is therefore vital, and such strategy typically encompasses data preparation (the initial cleaning and formatting); ongoing quality monitoring (detection of errors so that they can be corrected earlier rather than later); standardization (establishing enterprise-approved norms), enrichment (adding content or structure to otherwise incomplete information), and auditing (the formal recording and verification of the quality-control procedures). Data quality frameworks and tools are frequently used to automate and scale the analysis and rectification of data quality issues, especially for structured data sources. Data governance frameworks provide similar functionality for all aspects of data quality and are particularly essential to standardization, especially in complex organizations like healthcare payers.

7. Artificial Intelligence in Claims Automation

Artificial intelligence—especially natural language understanding, predictive modeling, and other machine-learning capabilities—has revolutionized various businesses and industrial sectors. However, the insurance industry has yet to realize widespread operational transformation with this powerful technology. The traditional claims-processing infrastructure remains largely unchanged, and the quality of insurance products and services has stagnated. Consequently, customer dissatisfaction is on the rise, and insurance payers are facing increasing regulation and scrutiny regarding fraud, waste, and abuse (FWA) in claims payments. The operational efficiencies and growing data assets associated with automation are increasingly being combined with artificial intelligence solutions to address these challenges. The focus is not only on handling claims automatically but also on augmenting human workflows with predictive capabilities to preempt and contain FWA, preparing for swift adaptation to changing regulations, assuring responsiveness to customer queries, and streamlining the broader customer journey.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Natural language understanding is being employed to automate claims adjudication by harnessing the wealth of unstructured data in encounter notes, clinician and provider comments and responses, and claim narratives. Predictive modeling informed by such data is being applied to recognize the most likely candidates for investigation and to assess the severity of any corrective actions. Given the historical failure of enterprise-wide end-to-end automation of insurance claims processing, emphasis is being placed on automating claims processing from insurer signals and data partnerships while handling more complex or nuanced claims through human-human collaboration. This is often realized through a human-in-the-loop framework, where AI augments human workflows, indicating next-best actions based on predictive models while leaving final decisions to trained human operators.

7.1. Natural language understanding for adjudication

Natural language understanding (NLU) is a subfield of natural language processing (NLP) focused on enabling autonomous systems to grasp human language with human-like comprehension and contextual understanding. When applied to adversarial claims processing, NLU models analyze the claim's narrative (interpretation) or reason posted by the provider and either assign it to a predefined class (e.g., cosmetic procedures) or determine whether to conduct a further specialist review. The models take advantage of the fact that such classes are often defined in policies, which usually exist in digital form and are structured for logical inference. Model interpretability is crucial, as misclassified claims can be mistaken either for permissible procedures that should receive payment or for impermissible claims that should not.

Deep-learning-based NLU agents for adjudication offer several advantages: they have been performing at or above human level across a spectrum of NLP tasks (such as sentiment classification), their performance improves with data size, and they can be fine-tuned using structured data. Moreover, they are potentially democratizing access to domain- and task-specific intelligence by making it accessible to non-experts and available as enterprise-grade services.

Claims-processing costs vary by payer and function's within-payer value chain, but most Divided function categories, severity categories, and within-category severity archetypes are common across payers. Their foundational assignment is thus ideally suited to prediction-and-classification solutions that are aware of context and afford inference at taxonomical and semantical levels.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Equation 3. Provider/pharma amount from total spend

Using the same 89% share and total spend $H = 3.8T$:

$$P = 0.89H$$

Substitute:

$P = 0.89 \times 3.8$ $P = 3.382$

Final form

$$\boxed{P = 3.382 \text{ trillion dollars}}$$

7.2. Predictive modeling for fraud, waste, and abuse

Concurrently, advances in explainable AI and predictive modeling techniques enable the automated detection of fraud, waste, and abuse (FWA) activity. For each claim, prior models consider an array of features—including historical patterns, member interactions, metric anomalies, peer group behaviors, and data quality aspects—to predict the probability of triangular abuse (e.g., billing for additional services that were neither rendered nor needed by the member), unqualified provider relationships (e.g., borderline rating but insufficiently recurrent), and the likelihood of a resulting forensic audit. All predictions are offered as natural language statements to be consumed by routine and responsive operational processes of the payer (e.g., to activate stop-watches for outbound assistants), and can be refined with elaborate post-hoc justification to support accountability. Supervision strategies rely on a periodically balanced mixture of parametric methods and non-labeled classification exploration.

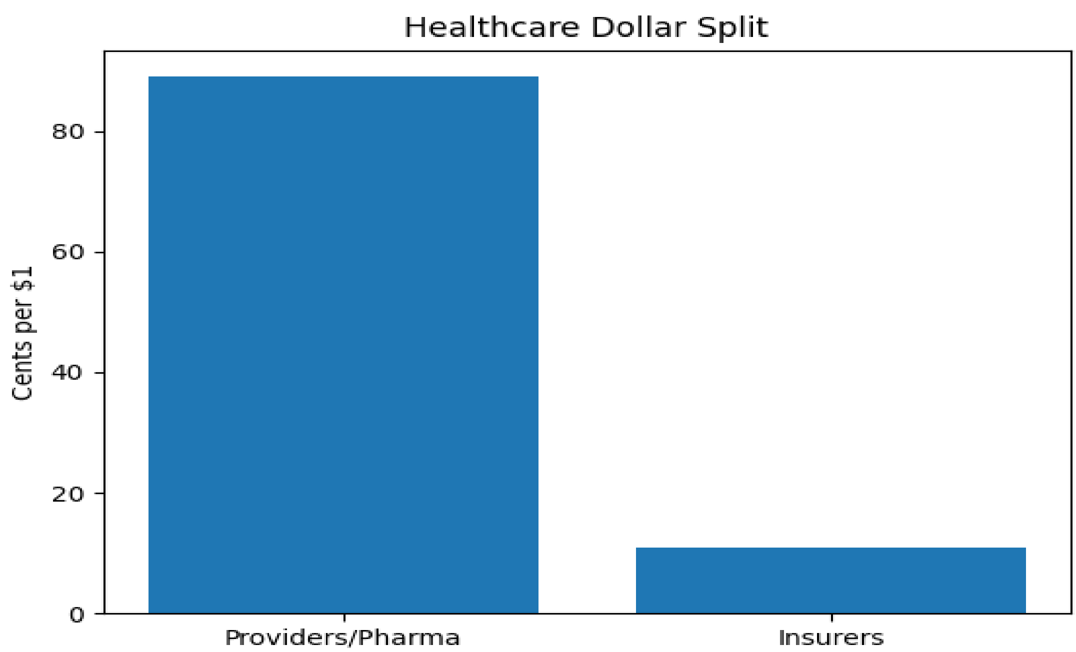
The detection of FWA and the prevention of administrative delays constitutes one side of the claims triangle. The other side relies on human oversight and business procedures to manage other potential sources of fraud, waste, and abuse not directly observable in real-time data but that require management intervention to be contained or disabled. Completeness of oversight is thus achieved by building predictive models for these secondary types of fraud, reduction, and abuse as well. Positioned at the core of those service extension components, predictive modeling applies advanced algorithms to all available data sources—transactions, services, and counterparty behavior—in order to forecast the presence of potential sources of risk.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07



8. Workflow Automation and Operations Transformation

Streamlining and transforming healthcare claim operations through automation. Until now, AI could only automate a portion of the claims lifecycle, requiring intervention from human staff on cases that could not be resolved through automated processing. Through natural language understanding for adjudication and predictive modeling for fraud, waste, and abuse, the system is undergoing evolution toward an end-to-end workflow that comprehensively addresses all aspects of the process, including data ingestion, validation, adjudication, fraud detection, and vendor management. Human effort on the exception cases needing domain expertise is now facilitated through a user-friendly, intelligent assistant that proactively recommends actions. Moreover, these actions are captured by the workflow automation layer that can learn from these interactions, thereby gradually building capability to address a larger share of claims.

Processing healthcare claims consumes significant resources for payers and impacts the economics of delivery systems and member experiences. End-to-end adoption of workflow automation, combined with AI-driven natural language understanding for adjudication and

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

predictive modeling for fraud, waste, and abuse, can significantly enhance efficiency. While these AI pillars have so far allowed automation of most parts of the workflow, they have required human participation in the small percentage of the claims needing domain expertise for decision-making. Realizing value from this effort has largely been limited to exception management. The move towards an intelligent assistant is now enabling automation of a large portion of that burden. By integrating full-cycle workflow automation with fraud-detection capabilities and state-of-the-art generative AI for claims processing, the goal is to deliver comprehensive intelligence-driven claims-processing automation for healthcare payers.

8.1. End-to-end claims lifecycle automation

Automating the entire claims lifecycle from claim submission to payments represents the primary opportunity for machine learning in healthcare claims operations. In this vision, AI-enabled true payer intelligence addresses aversion to infrastructure and technology specific to select phases of claims processing rather than the entire ecosystem. The vision integrates automated claims adjudication with fraud, waste, and abuse-fighting predictive models into a unified claims processing experience.

Such automation minimizes manual effort across the lifecycle of a claim—from initial registration through negotiation of complex provider contracts, to persistence through problems with data quality and sufficiency, to analysis of patterns and anomalies with historical claims, to claim closure and subsequent recovery of outstanding medical debts. Claims represent a natural area for the application of machine learning given that these advanced analytical techniques are most often deployed to perform repetitive work more quickly and accurately than humans, thereby freeing the latter to focus on more strategic tasks.

Fig 4: AI-Powered Claims Automation for Payers in Healthcare

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

8.2. Exception handling and human-in-the-loop

Exception handling presents unique challenges and opportunities for AI-enabled workflows. It is costly for a payer to process, and even a small percentage introduces a large financial burden. A specialized healthcare field relies on a highly trained, specialized team to adjudicate the complex claims. These small numbers also consume the majority of team members' time and require high degrees of expertise. AI enables infrequent exceptions to be predicted and a directed, specialized model to be designed for these edge-case claims. Focused models also improve expense ratios and leverage highly specialized workers efficiently.

Few such exceptions are reserved for human review. A human-in-the-loop capability monitors operationalizing AI models. Alerts canvas the region, checking model confidence as well as existing workflow and case management for additional supporting evidence and ground truth for the model. Thresholds fro the alerts are tuned for coverage and cost. The monitoring also maintains a strong knowledge base, learning from regular ups and downs in model accuracy as the data and domain changes.

9. Data Governance, Privacy, and Compliance

Data pipelines in healthcare are built on data sources with varying degrees of consistency and completeness. Health claims datasets, despite their size and breadth, are not consistent or inherently reliable. Privacy-preserving analytics, supported by differential privacy techniques, enables beneficial analytics on health claims datasets without compromising patient privacy. However, easily accessible health claims networks carry risks of leakage of sensitive information. Regulatory compliance is paramount. Not only should algorithms be defensible on regulatory requirements but every use of patient data for derivative applications should undergo strong auditing requirements.

Claims adjudication, even in a fully automated virtual payer, requires regulatory compliance, in particular support for timely deliveries and the maintenance of required quality of care standards. These requirements necessitate the introduction of a human-in-the-loop function capable of validating or invalidating unusual diagnoses and procedure combinations. Without such a support mechanism, such virtual payers will not satisfy the auditability requirement and will always be open to regulatory challenges in the near term.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 01 ISSUE: 01
RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Table 3. Automation-savings scenarios derived from the article

- processing cost = **8.9¢ per $1 healthcare spend**

Scenario	Formula	New processing cost
Baseline	8.9¢	8.90¢
15% reduction	8.9 × (1 − 0.15)	7.565¢
25% reduction	8.9 × (1 − 0.25)	6.675¢
30% reduction	8.9 × (1 − 0.30)	6.230¢

9.1. Privacy-preserving analytics

Limited access to training data is one of the greatest challenges in the development of machine learning models for healthcare payer operations. Due to the regulatory and technical hurdles associated with the sharing of sensitive healthcare data, applying machine learning to individual problems in a decoupled fashion, while maintaining data privacy, is often infeasible. Federal and state privacy regulations pertaining to the handling of insured patient data severely restrict access to training datasets. Even in the absence of regulatory restrictions, data owners are often reluctant to share sensitive data with third-party vendors — especially those engaged in predictive modeling for potential misuse. Enabling privacy-preserving machine learning through the use of synthetic data, federated learning, and differentially-private training holds the promise of alleviating data access concerns and accelerating the adoption of AI in healthcare payer operations.

Synthetic data has become a dominant approach for addressing data scarcity in the field of deep learning. Several generative adversarial networks and variational autoencoder architectures have been proposed for synthesizing medical image data, decision-tree-based tabular data, and longitudinal clinical data. A similar approach can be followed to enable claims-data synthesis for training claims-adjudication decision models. Synthetic claims data can be compactly modeled as a procedural document that mimics the cadence and composition of a JSON object but stores each attribute, such as the claim line items, leukemias in ludicrously rich detail.

9.2. Regulatory considerations and auditability

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Regulatory compliance and auditability must be holistically incorporated into processes when introducing automation capabilities. Regulations in the functions supported by automation must be examined to ensure the compliance capabilities not only meet existing regulatory requirements but proactively and preemptively accommodate potential future regulations.

Governance policies for automated functions—basically the same as for any automated decision-support function—must be established. An understanding of the reasons for ratios of true positives to false positives is necessary, including a view on the quantification and cost benefit of addressing false positives/negatives as exceptions. Input into decisions on free-text fields for automated claims approval must derive from full knowledge of past patterns and trends, along with specifics of the automated recognition process. Rules-based conditions requiring human intervention need to be reviewed to ensure both their necessity and continued relevance. Self-learning capabilities on such exceptions must be viewable in order to determine whether they could now be incorporated into the automation process.

Equation 4. Insurer-retained amount from the 11% rule

$$I = 0.11H$$

Substitute $H = 3.8$:

$I = 0.11 \times 3.8\ I = 0.418$

Thus

$$\boxed{I = 0.418 \text{ trillion dollars} = 418 \text{ billion dollars}}$$

Important consistency note

The insurers are "costing them nearly $42 billion annually," but mathematically,

$$0.11 \times 3.8T = 418B$$

not $42B. So either:

- the 11% figure is being used differently,
- or the $42B refers to a narrower cost category,

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

- or there is a typo in the article.

10. Value Creation and Economic Impacts

Accelerated pipeline and operations automation would create substantial value and economic impacts across healthcare payer organizations. With technology and vendor readiness in place, functions and workflows of a complete department could realize breakthrough efficiency gains, and entire organizations could derive substantial expense savings while addressing claims-capture and other operational challenges through AI-enabled end-to-end lifecycle automation. Cost reductions through automation would not only drive higher profit margins but also improve the quality of care delivered to members and the overall experience of interacting with the payer.

Evidence from nonhealthcare industries suggests that AI-powered automation implementations can reduce costs by 25% or more while driving 40%–60% increases in throughput and substantial improvements in quality and customer satisfaction. Similar results would be expected from deployments in healthcare payers, and several cloud and native AI technology vendors are building claims-and-operation automation solutions capable of realizing many of these advantages.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

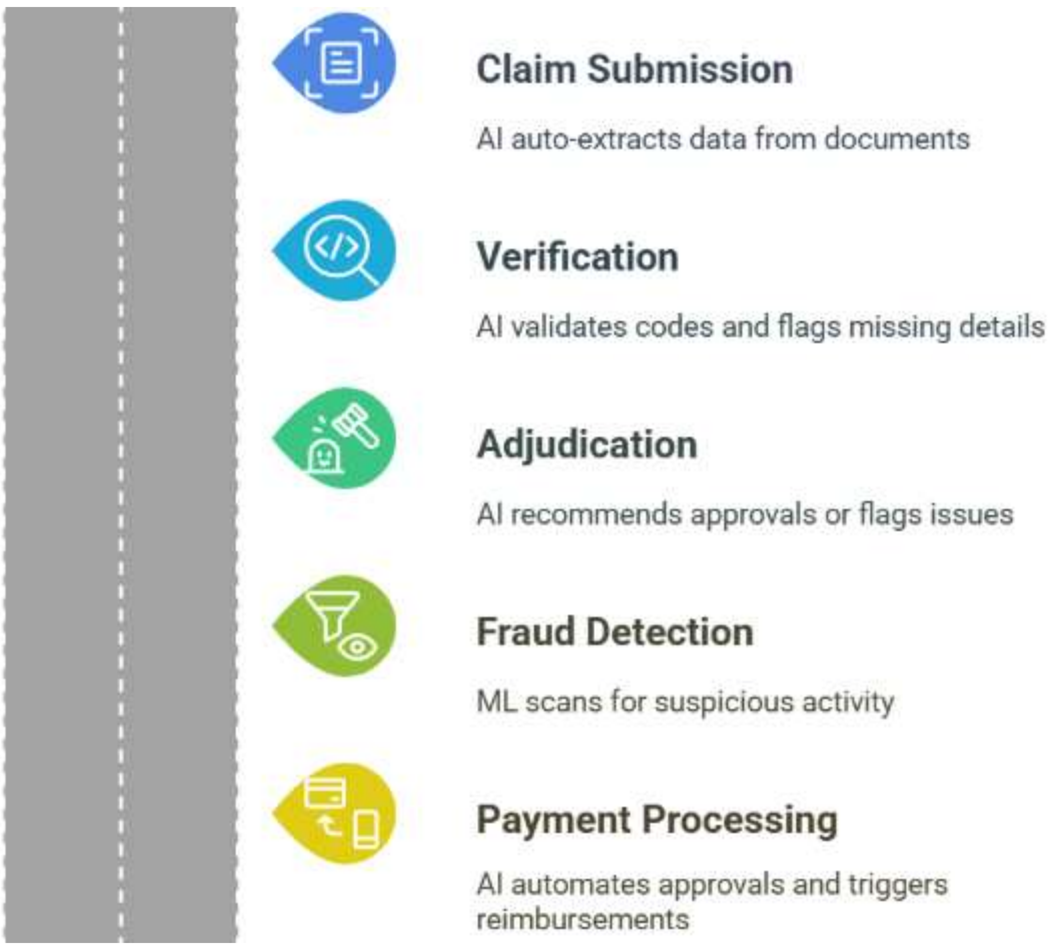


Fig 5: AI-Powered Claims Processing Journey

10.1. Cost reduction and efficiency gains

Healthcare payers face mounting pressure to deliver affordable care while also becoming more efficient. Accelerating the Digital Transformation of Payer Operations Payer organizations have begun investments in artificial intelligence to reduce operating costs, with business process outsourcing providers further promoting investments in digital labor and intelligent automation capabilities.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Payer case studies point to 15%–30% reductions in operating costs directly related to claims automation. Notably, however, these savings estimates do not take into account the full potential of their investments. End-to-end automation of the claims lifecycle, predictive modeling for fraud, waste, and abuse detection, and a lens into the quality of third-party services can also unlock significant additional savings and efficiency gains—a reality that leading-edge payers now recognize as their operations become AI-enabled.

10.2. Quality of care and member experience

Increased operational efficiency has a cascading effect on quality of care, as improved speed to decision for coverage and reimbursement results in expedited claims payments for healthcare providers, reducing financial stress. Moreover, deliverable assurances of accurate and complete payments minimize post-payment clawbacks and audit hassle for providers. By relieving health plan backlogs, automated exception handling augurs well for both the quality of service from health plans to healthcare providers, and from healthcare providers to members.

In addition, automating the end-to-end claims lifecycle facilitates seamless integration with complementary functions. For instance, machine-generated summaries for manually adjudicated claims can be enriched with the findings of concurrent predictive fraud models, supporting prompt communications to affected members. Within parts of the organization where full automation is not yet possible, integration with human-in-the-loop systems ensures that the quality of service suffers as little as possible. Overall, organizational transformations diminish the burden of claims processing, restore organizational focus, and free up people to work on areas that require a human touch, such as member outreach and engagement.

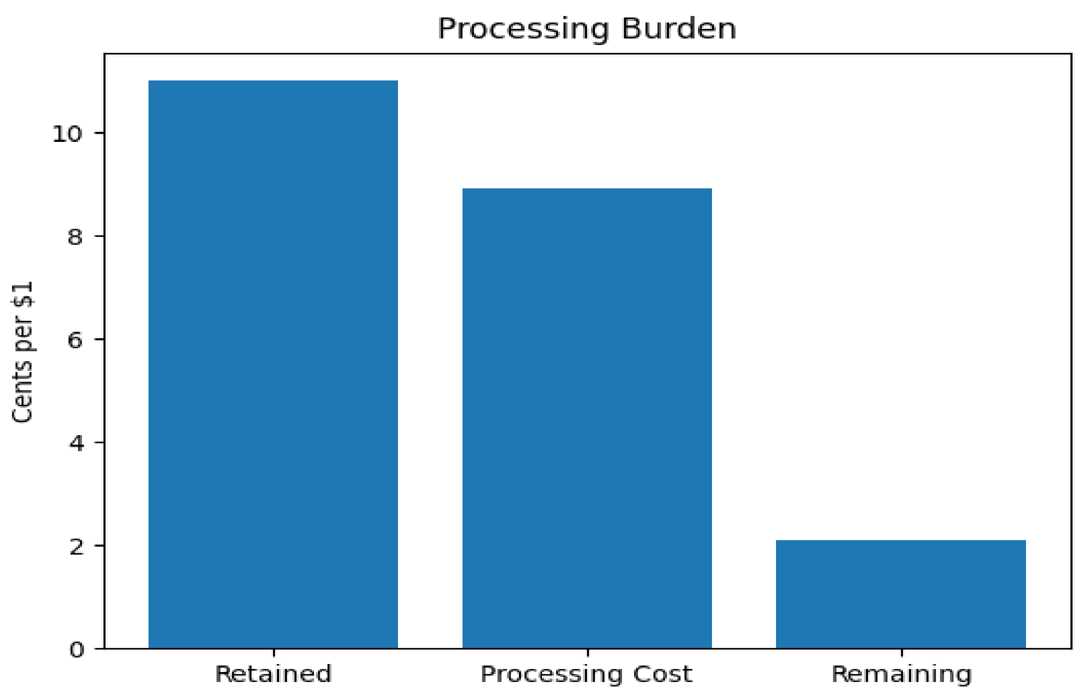
11. Implementation Considerations and Roadmaps

Successful execution of large-scale AI programs requires careful deliberation of three principle considerations. First, aspiring leaders must assess their AI capabilities in conjunction with their strategic priorities. Concurrently, they must evaluate the technology solution landscape, informing the final selection. And ideally, such evaluation should incorporate vendor offerings and the broader ecosystem. A closer look at these areas lends insight into how AI can best enable business imperatives, and de-risk the corresponding investments.

As AI applications multiply within the sector, the technology landscape is also evolving rapidly; now, a wider set of both standard and bespoke AI tools is commercially available. Unsurprisingly, the specific technology choices—including whether to build, buy, or partner—

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 01 ISSUE: 01
RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

remain vital elements of all individual AI initiatives. For larger-scale change—such as end-to-end claims-lifecycle automation—criteria for such decisions should exist not only at the application level but also in terms of overall architecture. Broadly, they can include considerations about whether to source a suite of AI tools from a single provider or integrate capabilities from multiple vendors.



Equation 5. Claims-processing expenditure equation

Claims processing expenditures can reach at least 10% of the claims' payment value.

Let:

- C_p = claims-processing expenditure
- V_c = claims payment value

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 01 ISSUE: 01
RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Then the minimum relationship is

$$C_p \geq 0.10 V_c$$

If we interpret claims payment value as the 89% going to providers/pharma, then

$$V_c = 0.89H$$

So

$C_p \geq 0.10(0.89H)$ $C_p \geq 0.089H$

Substitute $H = 3.8$:

$C_p \geq 0.089 \times 3.8$ $C_p \geq 0.3382$

Thus,

$$\boxed{C_p \geq 0.3382 \text{ trillion dollars} = 338.2 \text{ billion dollars}}$$

Per-dollar interpretation

For each $1 of healthcare spend:

$$C_p \geq 0.10 \times 0.89 = 0.089$$

So

$$\boxed{C_p \geq 8.9 \text{ cents per healthcare dollar}}$$

11.1. Readiness assessment and strategic alignment

As with any major strategic initiative, organizations considering data- or AI-led automation should assess their capabilities, resources, and infrastructure to understand readiness and define execution scope. Questions to consider include the follow-ing: Have analytics ever been applied to claims? If so, are the results still relevant? What operational areas have historically been more or less automated? Are all procedures well documented, and is knowledge easily transferable? What areas are particularly sensitive to poor data quality? Will soon-to-be-

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

required documentation provide a level of detail or consistency enabling data-to-AI implementations? Are manual and automated sub-processes equally competent? Are past vendor solutions, considered "black-box," receiving renewed interest? What's the relationship with AI vendors? Is local data center capacity and skill sufficient for DataOps management, given recent cloud-outside-a-cloud discussions? Are audit support solutions in place? What are the pressures to reduce costs—AI-generated solutions, shrinking budgets, or union actions?

Strategic alignment is another crucial element as investment and operational priorities must be approved by management. Roadmaps mapping state-changes against expected AI-enabled impact and experience can ensure that the data pipelines address and prioritize current concerns. AI vendor expenses for labs reflecting future capability should be evaluated. The importance of look-alike analysis can partly determine how soon DataOps teams will need vendor assistance. Some business areas may require more immediate support, such as productivity, solution testing, or better quality offers/templates. Investing in data-to-AI pipelines may proceed based on these considerations.

11.2. Technology selection and vendor landscape

Decision-making toolkits are needed to compare potential approaches across the relevant dimensions and to help organizations select the most appropriate option for their context. Vendor-provided automation technologies should be carefully evaluated relative to the unique requirements of the problem at hand. If user-defined, vendor-provided solutions are selected, the quality of the vendor's natural language processing, machine learning, optical character recognition, and data integration technology must be assessed. Vendors providing specific functionality should be assessed for their partnerships with complementary technology companies.

Cloud-based technology stacks from hyperscalers provide strategic advantages in terms of integrated technology, security, scalability, performance, and cost efficiencies afforded through a pay-as-you-scale model. With multiple cloud providers offering compatible platforms, organizations may choose to build on a single provider's platform or combine the complementary advantages of multiple providers. Such service-level agreements ensure that regulators can obtain the necessary level of oversight without compromising privacy. The scope for service-level agreements to define and regulate these operations is in its early stages of development.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 01 ISSUE: 01
RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

Table 4. Throughput-improvement scenarios

Let baseline throughput be T_0.

Improvement	Formula	New throughput
40% increase	$T = T_0(1 + 0.40)$	$1.4T_0$
50% increase	$T = T_0(1 + 0.50)$	$1.5T_0$
60% increase	$T = T_0(1 + 0.60)$	$1.6T_0$

12. Results

The following section reports AI-based automation efforts directly undertaken by the healthcare payer community. While many such implementations were initiated or heavily supported by AI technology companies, the organization overseeing deployment typically remains an insurance company.

An early production-grade deployment aims to eliminate medical coding operations during claims processing phases such as pre- and post-adjudication. Natural Language Processing (NLP) transforms unstructured provider documentation into structured records. These documents have reduced complexity for machine learning models, enabling a modestly supervised framework requiring as few as 100 images. Deploying the NLP-generated procedures and diagnoses directly into an AI-adjudicated workflow has cut coding times by half within a claim from prior copays and almost completely truncated subsequent coding activities. Both the size of second-pass coding teams and overall turn-time remain under pressure.

Another adoption claims to automate the whole lifecycle of care management. Incorporating several components, Clinical AI has reduced manual efforts by fraud detection, social determinants gathering, and patient journey modelling. The last component generates optimal treatment pathways, risk scores, and predictive models filtering patients by support modalities. These AI techniques inform Managed Care Organization interventions across an annually rising membership sourced not only from its parent Payer but also from external allies. Recently recorded results include 37% improvement in recurrent breast cancer identification and 90% alteration in chest sticking therapy provisions. Beyond successful business impact–counting over 330 million over the past five years–the solution is set to fold growth trajectory and optimizations within Clinical Services portfolios.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

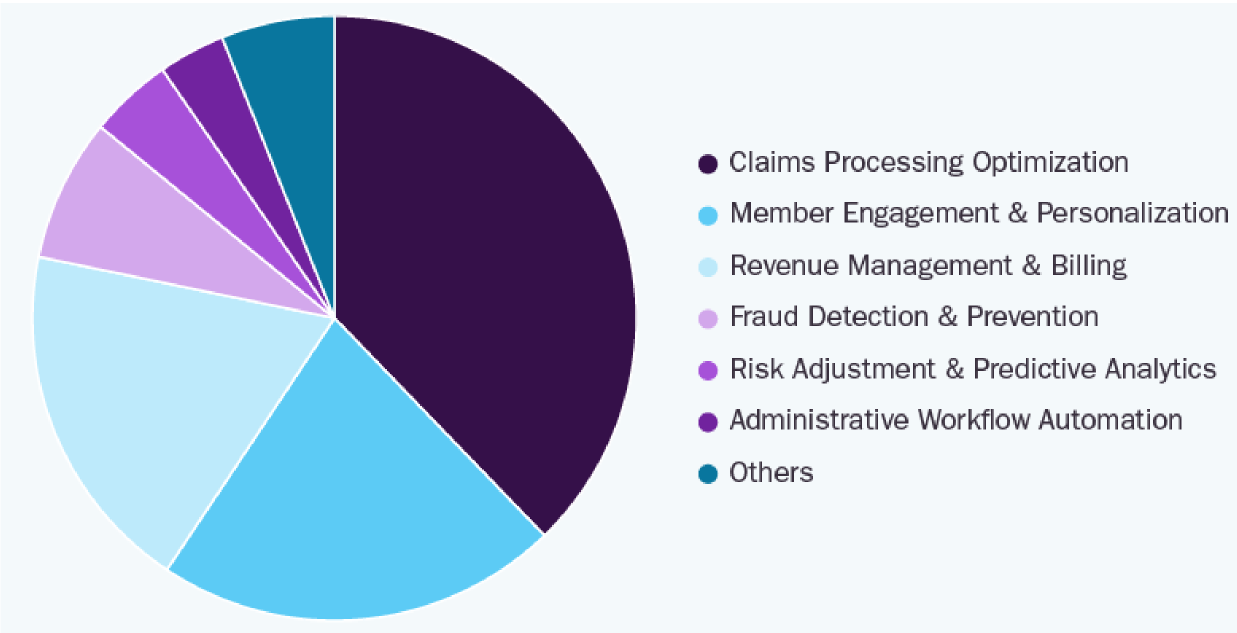


Fig 6: Artificial Intelligence For Healthcare Payer Market Report

12.1. Payer-led automation implementations

Automation has transformed processes across many industries; however, healthcare payers have yet to leverage that power to improve operational efficiency and effectiveness. AI-enabled payer intelligence, the use of AI-powered automation in payer operations, addresses this gap by streamlining and automating claims processing, enabling providers to allocate revenue toward improving the quality of care rather than processing administrative transactions. These efforts deploy product logical automation: natural language understanding algorithms process unstructured narratives, predictive models for fraud, waste, and abuse reliance determine further investigation needs, and surpluses or deficiencies are routed for manual review.

Automation is also being applied to claims lifecycles beyond initial adjudication. Product logical automation is already being integrated into the claims lifecycle decisions and processes, replacing workers on parts of the decision tree that are relatively simple and predictable with low probability of exception handling. Investments in such area of claims processing are expected to grow rapidly in the coming years, and enterprise-wide operational infrastructure services are

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

starting to embrace product logical automation. Automated decisions considered to require human judgment are now being sent to workers who independently process cases routed to them.

13. Conclusion

Healthcare payers face challenges in data volume and claims processing that can be alleviated through AI. Transparent, auditable, and privacy-preserving data pipelines enable retrospective and prospective claims analysis. Automated Natural Language Processing of claim files allows for automated adjudication based on institutional knowledge and predictive modeling based detection of Fraud, Waste, and Abuse. Fully automating the claims life cycle provides vast efficiency opportunities, with Human-in-the-Loop AI techniques enabling exception handling. The described constructs fulfill Gartner's Payer Intelligence category.

Despite the many inefficiencies in claims processing and servicing that third-party payers incur, particularly in the Manual and Human-in-the-Loop areas, payers have so far not undertaken a concerted effort to automate many of these processes. As the pressure on costs mounts, a number of these workflows are being automated using data pipelines, privacy-preserving techniques, Natural Language Processing and other Artificial Intelligence methods, thus meeting a number of Gartner's criteria for Payer Intelligence. The technologies are mature and available in the market today, and a number of organizations have pioneered implementations. Organizations not currently pursuing such opportunities may wish to assess their readiness to do so in the near future.

References

1. Abdel-Basset, M., Mohamed, R., Elhoseny, M., Chakrabortty, R. K., & Ryan, M. J. (2020). A hybrid COVID-19 detection model using an integrated framework of deep learning and swarm intelligence. International Journal of Intelligent Systems, 35(12), 1899–1920.
2. Ahmad, T., Zhang, D., Huang, C., Zhang, H., Dai, N., Song, Y., & Chen, H. (2021). Artificial intelligence in sustainable energy industry: Status quo, challenges and opportunities. Journal of Cleaner Production, 289, 125834.
3. Albahri, A. S., Hamid, R. A., Alwan, J. K., Al-Qaysi, Z. T., Zaidan, A. A., Zaidan, B. B., Albahri, O. S., Alamoodi, A. H., & Salih, M. M. (2021). Role of biological data mining and machine learning techniques in detecting and diagnosing diseases: A systematic review. Artificial Intelligence Review, 54(5), 3473–3513.

4. Alouffi, B., Hasnain, M., Alharbi, A., Alosaimi, W., Alyami, H., & Ayaz, M. (2021). A systematic literature review on cloud computing security: Threats and mitigation strategies. IEEE Access, 9, 57792–57807.
5. Pereira, K., Vinagre, J., Alonso, A. N., Coelho, F., & Carvalho, M. (2022, September). Privacy-preserving machine learning in life insurance risk prediction. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 44-52). Cham: Springer Nature Switzerland.
6. Bousdekis, A., Apostolou, D., & Mentzas, G. (2021). A review of data-driven decision-making methods for industry. Information Systems Frontiers, 23(5), 1151–1171.
7. Davenport, T. H., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. Journal of the Academy of Marketing Science, 48(1), 24–42.
8. Mattaparthi, R. (2023). Connected Fleet Intelligence: Edge-Centric Analytics and Computer Vision for Predictive Manufacturing and Asset Resilience. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 6(5), 9077-9088.
9. Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., … Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994.
10. Elgendy, N., & Elragal, A. (2020). Big data analytics in support of the decision-making process: A systematic literature review. Journal of Decision Systems, 29(2), 138–160.
11. Gandomi, A., & Haider, M. (2020). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 51, 102040.
12. Kolla, S. K. (2023). Learning Health Systems Machine Intelligence for Clinical Prediction and Healthcare Optimization. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 6(2), 7955-7966.
13. Gupta, M., Abdelsalam, M., Khorsandroo, S., & Mittal, S. (2020). Security and privacy in smart healthcare systems using artificial intelligence: A survey. ACM Computing Surveys, 53(6), 1–36.
14. Janssen, M., van der Voort, H., & Wahyudi, A. (2020). Factors influencing big data decision-making quality. Journal of Business Research, 110, 338–345.
15. Reddy, V. A. R. (2022). Data-Driven Healthcare Operations: Architecting Unified Member, Provider, and Claims Intelligence Platforms. International Journal of Science, Research and Technology, 5(5), 8511-8521.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

16. Kaur, D., Uslu, S., Rittichier, K. J., & Durresi, A. (2022). Trustworthy artificial intelligence: A review. ACM Computing Surveys, 55(2), 1–38.
17. Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., Shiraz, M., & Gani, A. (2021). Big data: Survey, technologies, opportunities, and challenges. The Scientific World Journal, 2021, 712826.
18. Aitha, A. R. (2023). Cloud-Native Big Data AI/ML Framework for Risk Intelligence and Fraud Control in Banking and Insurance Ecosystems. Available at SSRN 6157967.
19. Liang, H., Tsui, B. Y., Ni, H., Valentim, C. C. S., Baxter, S. L., Liu, G., Cai, W., Kermany, D. S., Sun, X., Chen, J., He, L., & Zhang, K. (2020). Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nature Medicine, 26(5), 699–704.
20. Lu, Y. (2020). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 7(1), 1–29.
21. Mangala, N. (2022). Implementing Databricks Unity Catalog For Centralized Data Governance In Multi-Business-Unitenterprises. Journal of International Crisis and Risk Communication Research, 101-122.
22. Rai, A., Constantinides, P., & Sarker, S. (2021). Next-generation digital platforms: Toward human–AI hybrids. MIS Quarterly, 45(1), iii–ix.
23. Radanliev, P., De Roure, D., Walton, R., Van Kleek, M., Santos, O., Ani, U., & Montalvo, R. M. (2020). Artificial intelligence and cyber risk in critical infrastructures. AI & Society, 35(4), 1221–1232.
24. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2021). Organizational decision-making structures in the age of artificial intelligence. California Management Review, 63(3), 66–83.
25. Yandamuri, U. S. (2023). An Intelligent Analytics Framework Combining Big Data and Machine Learning for Business Forecasting. International Journal Of Finance, 36(6), 682-706.
26. Topol, E. J. (2021). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 27(1), 44–56.
27. Wamba, S. F., Queiroz, M. M., Wu, L., & Sivarajah, U. (2020). Big data analytics-enabled sensing capability and organizational outcomes: Assessing the mediating effects of business analytics capability. Journal of Business Research, 113, 180–194.
28. Aggarwal, C. C. (2023). Data mining for business analytics and decision support. ACM SIGKDD Explorations Newsletter, 25(1), 1–12.
29. Albahri, O. S., Albahri, A. S., Mohammed, K. I., Zaidan, A. A., Zaidan, B. B., Hashim, M., & Salman, O. H. (2022). Systematic review of artificial intelligence techniques in healthcare data analytics. Artificial Intelligence Review, 55(5), 3917–3975.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

30. Kolla, T., & Kolla, S. K. (2023). FHIR-Based Real-Time Healthcare Analytics using Unsupervised Learning. International Journal of Future Innovative Science and Technology (IJFIST), 6(6), 11751.
31. Ali, O., Abdelbaky, M. A., Shrestha, A., Elbasi, E., Alryalat, M. A. A., & Dwivedi, Y. K. (2023). A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. Journal of Innovation & Knowledge, 8(1), 100333.
32. Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2022). Software engineering for machine learning: A case study. Communications of the ACM, 65(7), 56–65.
33. Gottimukkala, V. R. R. (2020). Energy-Efficient Design Patterns for Large-Scale Banking Applications Deployed on AWS Cloud. power, 9(12).
34. Araci, D. (2022). FinBERT: Financial sentiment analysis with pre-trained language models. Applied Artificial Intelligence, 36(1), 203–219.
35. Batarseh, F. A., & Freeman, L. (2022). AI assurance: Toward trustworthy artificial intelligence in enterprise systems. IEEE Intelligent Systems, 37(5), 76–84.
36. Bandi, V. D. V. K. (2023). MLOps frameworks for reliable model deployment in cloud data platforms. Journal of Artificial Intelligence and Big Data, 3(1), 84-101.
37. Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2021). Artificial intelligence in organizations: Current state and future opportunities. MIS Quarterly Executive, 20(4), 241–263.
38. Bertsimas, D., Delarue, A., Mullainathan, S., & Obermeyer, Z. (2022). From predictive to prescriptive analytics. Management Science, 68(1), 102–131.
39. Chui, M., Hall, B., Mayhew, H., & Singla, A. (2022). The state of AI in 2022—and a half decade in review. McKinsey Quarterly, 3, 1–16.
40. Inala, R. Designing Scalable Technology Architectures for Customer Data in Group Insurance and Investment Platforms.
41. Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2022). Big data in healthcare: Management, analysis, and future prospects. Journal of Big Data, 9(1), 54.
42. Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, C., Jebabli, I., … Wirtz, J. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI. International Journal of Information Management, 71, 102642.
43. Ebert, C., & Louridas, P. (2023). Generative AI for software engineering. IEEE Software, 40(4), 30–38.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

44. Guo, Y., Li, Y., & Liu, X. (2022). Machine learning-based intelligent claims processing in insurance: A review. Expert Systems with Applications, 198, 116756.
45. Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
46. He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2022). The practical implementation of artificial intelligence technologies in medicine. Nature Medicine, 28(1), 31–38.
47. Huyen, C. (2022). Designing machine learning systems: An iterative process for production-ready applications. O'Reilly Media.
48. Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2023). Data governance for artificial intelligence: Frameworks and challenges. Government Information Quarterly, 40(1), 101728.
49. Nagabhyru, K. C., & Engineer, S. D. (2023). Unifying Data Engineering and Machine Learning Pipelines: An Enterprise Roadmap to Automated Model Deployment.
50. Kaul, V., Enslin, S., & Gross, S. A. (2023). History of artificial intelligence in medicine. Gastrointestinal Endoscopy, 97(4), 807–812.
51. Kshetri, N. (2022). Artificial intelligence in insurance and claims management: Trends and implications. IT Professional, 24(5), 20–27.
52. Sarker, I. H. (2022). AI-based modeling: Techniques, applications, and research issues. SN Computer Science, 3(2), 158.
53. Wang, L., Alexander, C. A., & Pan, Y. (2023). Artificial intelligence in insurance: Applications, challenges, and future directions. IEEE Access, 11, 42135–42152.
54. Ahmed, E., Yaqoob, I., Gani, A., Imran, M., & Guizani, M. (2020). Internet-of-Things-based smart environments: State of the art, taxonomy, and open research challenges. IEEE Wireless Communications, 27(5), 10–16.
55. Inala, R. Advancing Group Insurance Solutions Through Ai-Enhanced Technology Architectures And Big Data Insights.
56. Albahri, O. S., Zaidan, A. A., Albahri, A. S., Zaidan, B. B., Hashim, M., Salman, O. H., & Abdulkareem, K. H. (2023). Artificial intelligence techniques for decision support systems in healthcare and insurance: A systematic review. Artificial Intelligence Review, 56(8), 8139–8181.
57. Baesens, B., Höppner, S., Verdonck, T., & Verbeke, W. (2021). Machine learning techniques for fraud detection and claims analytics: A review. Decision Support Systems, 150, 113523.
58. Davuluri, P. N. AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

59. Bhardwaj, A., Garg, A., Ram, S., Gaur, V., & Koundal, D. (2022). Cloud computing-based intelligent data management using artificial intelligence: A review. Journal of King Saud University – Computer and Information Sciences, 34(8), 5367–5381.
60. Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work. National Bureau of Economic Research Working Paper No. 31161.
61. Cao, L. (2022). AI in finance: Challenges, techniques, and opportunities. ACM Computing Surveys, 55(3), 1–38.
62. Mangalampalli, B. M. (2022). Automated Invoice Validation Systems Using Advanced SQL Analytics in Healthcare Insurance. Front Health Inform, 11.
63. Chen, M., Mao, S., Zhang, Y., & Leung, V. C. M. (2020). Big data: Related technologies, challenges, and future prospects. Springer.
64. Davenport, T. H., & Mittal, N. (2022). All-in on AI: How smart companies win big with artificial intelligence. Harvard Business Review Press.
65. Dirican, C. (2020). The impacts of robotics, artificial intelligence on business and economics. Procedia - Social and Behavioral Sciences, 195, 564–573.
66. Dubey, R., Bryde, D. J., Foropon, C., Tiwari, M., Gunasekaran, A., & Papadopoulos, T. (2021). Artificial intelligence for improving operational excellence in organizations. International Journal of Production Research, 59(18), 5603–5623.
67. Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2020). AI in business information systems. Business & Information Systems Engineering, 62(4), 277–281.
68. Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in healthcare. The Lancet Digital Health, 3(11), e745–e750.
69. Jarrahi, M. H. (2023). Artificial intelligence and the future of work: Human–AI collaboration in knowledge-intensive tasks. Business Horizons, 66(2), 181–193.
70. Jordan, M. I., & Mitchell, T. M. (2020). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255–260.
71. Kotu, V., & Deshpande, B. (2021). Data science: Concepts and practice (2nd ed.). Morgan Kaufmann.
72. Lee, J., Davari, H., Singh, J., & Pandhare, V. (2021). Industrial AI and predictive analytics for intelligent enterprises. Manufacturing Letters, 18, 20–23.
73. Ng, A. (2021). MLOps: From machine learning to production systems. Communications of the ACM, 64(12), 44–51.
74. Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 01 ISSUE: 01

RECEIVED: OCTOBER 25 REVISED: NOVEMBER 04 ACCEPTED: NOVEMBER 28 PUBLISHED: DECEMBER 07

75. Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3), 160.
76. Topol, E. J. (2023). Deep medicine: How artificial intelligence can make healthcare human again (Updated ed.). Basic Books.