



Original Article

AI-Driven Reusable Unified Extract for Multi-State Medicaid and Federal Reporting- a Product that saves Millions of Taxpayer Money through process efficiency and reusability

Mani Kanta Pothuri
Independent Researcher, USA.

Received On: 13/10/2025 **Revised On:** 14/11/2025 **Accepted On:** 27/11/2025 **Published On:** 07/12/2025

Abstract: The state and federal government policy developers of the United States are continuously exploring ways to enhance the strength of administering Medicaid facilities. Artificial intelligence (AI) is under research for implementation to offer these facilities to stakeholders, based on insights into security and the benefits of integrated data management. These processes involve different patient data dimensions, including insurance, medical history, diagnostic records, and other electronic healthcare records. The services. This technology enhances capabilities to support healthcare processes through AI-driven content utilization frameworks. These allow optimization of MEDICAID content extraction at the national level, with integrity, to create esteem extracts/ reports. Increasing the automation of complex content flows and enabling effective use of the content across various functions. This paper presents a literature review and reports on an AI-powered, reusable data processing framework for generating accurate reports on Medicaid services at the state and federal levels. This follows recommendations to increase processing esteem through these advancements.

Keywords: MEDICAID, AI, Automation, Reusable data processing, Integrated data management.

1. Introduction

Medicaid support provisions across states and at the federal level require continuous data exchange and processing content. This is extracted for ICN, ACHN, PERM, MMIS, and other approvals, based on the tenure involved in processing claims. The latest budget reports underscore the importance of reconciling the regulations and legal guidelines governing human resources for Medicaid processing expenses. These highlight the importance of implementing significant policies and system modifications to support ongoing eligibility verifications and reduce federal government expenses for processing this content. The advanced needs and alignment with national and state-level governing policies position AI as a critical technology in the eyes of experts. This could be further streamlined using reusable data study frameworks for cost-effective and rapid data processing. This modernization supports states in handling such demands effectively and operating to enhance productivity, making administrative activities effective, [2]. AI is reducing staffing needs and the time required for resource management. Using continuous, high-speed network connectivity increases the efficiency of administrative processes and optimally serves American citizens eligible for healthcare support. The new federal policy requirements for Medicaid budget allocations for provision are considering AI mechanisms and different levels of technical sophistication. These are anticipated to support

states in handling such demands and in optimizing procedures to make administrative processes more effective. This reduces staffing and time intensity for content management and analysis. Utilizing resources effectively requires exploring available options, such as developing an AI data reuse framework to ease use and protect stakeholder needs. Using advancements could be an impact-driven approach in place of a simple bill on AI use, as implementing appropriate admin policies would have various downstream impacts.

Reusing the data in this format helps address the processing complexities that arise with the implementation of AI. In these ways, Medicaid processes are implementing policy and program variations to prepare for reforms. These modifications need advanced resources. This is suitable for MEDICAID regulations and eligibility criteria for clam processing. Policymakers across different states are now undergoing such modifications in view of the emerging and complex healthcare landscape and are embarking on budgeting and risk management. The decreased federal monetary reserves against policy funds are also compelling the use of AI data structures for verifying eligibility and handling emerging challenges. The regulations also included a targeted financial budget of \$200 million in 2026 to empower states or meet distinctive operational needs. These also involved allocating \$50 billion to healthcare projects

and supporting policy implementation in line with the latest changes across the financial years 2026 and 2030.

Different Medicaid processing units at various state levels require managing finances to cover associated administrative and program-based costs fully. These are resulting in enhanced stress on states to execute programs effectively [4]. Advanced requirements may accelerate efforts and improve the efficiency of new technologies, enabling easier handling of documentation processes by setting up eligibility authentication processes and claims.

Privately managed healthcare systems are leveraging advancements to reduce administrative stress and automate processes to avoid redundancies. The use of digital advancements is regarded as a prominent element of the plan to increase the efficiency of policy-based care. Reusing data architecture enhances innovation capabilities and healthcare technology infrastructure [5]. These could be further empowered by using an income assessment application for citizens based on eligibility. This follows connecting with payroll and economic data efficiently. These allow confirmation of eligibility for paperwork approvals. Natural language processing could be used to address the accessibility of Medicaid services for various users. Natural language processing also emerges as critical for assessing accessibility and enhancing proficiency to address

disabilities, as well as digital intelligence disparities while managing claims. Administrative overheads for executing programs like Medicaid could be the highest and increase complications [6]. The estimated values for these expenses may be about 15% to 30% of the government's complete care provision expenses.

2. AI-motivated data acquisition and integration architecture.

Automation processes involve a set of guidelines for accomplishing activities in a consistent manner. This could be enabled by learning algorithms that leverage content to detect patterns. These provide suggestions for new choices based on previous data and instances. Implementing Gen AI effectively allows the development of advanced content in various formats aligned with learning trends. Extraction of content using AI transforms business processes by leveraging details, as technology implements AI to detect and organize valuable content from different resources. This implements an effectively structured mechanism that involves users. These AI tools for extracting content and medical reports from different resources include direct optimization. These allow identification and interpretation of content with maximized accuracy.

Table 1: End-to-End AI-Enabled Data Extraction and Integration Framework

Component	Key Actions	Outcomes
Data Architecture & Design Standards	Enterprise-wide governance, consistent naming, schema definitions, metadata structures	Interoperability, CMS/federal compliance
Data Modeling & Semantic Layer Integration	Unified data model, semantic/cognitive layer	Automation, traceability, reporting alignment
Design Extract Templates & Data Formats	Reusable templates (CSV, XML, JSON, EDI), file naming, validation, schedules	Standardized reporting, compliance-ready extracts
Build Custom Extraction Scenarios	Dynamic pipelines (ICN, PERM, MMIS), business rules, metadata mappings	Precision, adaptability across states
Business Rule Automation	AI-driven validation, anomaly detection, correction workflows	Reduced errors, less manual intervention
AI Integration for Data Flow Optimization	ML routing paths, dependency scheduling, predictive extract patterns	Efficiency gains, improved delivery accuracy
Testing and Validation Framework	Automated unit/regression testing, anomaly detection, reconciliation dashboards	Quality assurance, compliance transparency
Security, Encryption, and Compliance	AES-256/FIPS encryption, RBAC, secure APIs, audit trails	HIPAA/CMS compliance, secure data exchange
AI-Based File Delivery and Monitoring	Automated scheduling/delivery, real-time dashboards, proactive error handling	SLA compliance, reliable transmissions
Scalable Deployment Framework	Modular plug-and-play, customizable components	Multi-state scalability, rapid adoption

This is the initial stage to implement AI automating paradigm in Medicaid service-delivering activities. This needs to be set up in accordance with appropriate governance standards. These include establishing an ethical framework as critical for Medicaid processes. These support the development and investment in technical advancements. These allow for mitigating negativity. The governing authorities need to align on transparency and mitigate biases to remediate AI-related issues for various stakeholders. The algorithm needs to ensure transparency in AI logic, exclude

biases, and develop informed outcomes. The data aspects must follow the required storage and naming standards. These follow the determining policy management scheme for Medicaid and promote maximum interoperability for various departments, as well as guidelines.

3. Creating a data model by unifying with semantics

The development of an integrated data model involves the use of AI's transformation capabilities for content extraction and the prominent management of activities, combined with a wide range of processing and automation mechanisms. This enhances automation capabilities and maximizes tracking. These follow developing reports and align with required standards. Executing unified approaches helps manage process time, reducing it by 40 to 60%.

3.1. Formatted extraction

Using AI increases cost efficiency as compared to human analysis processes. These allow developing extraction formats, which critically include creating a process map to identify information, creating extraction elements, and checking different documentation types for configuration and

setting up exception management processes [8]. These could be used as integrated procedures to depict process flows and policies to support the efficient development of appointment timelines. These obtain information about available dates and are followed by automating the processes.

3.2. Tailored data formatting

The data processing pipelines implemented for these processes include IC, PERM, and various report creation scenarios. These involve creating and managing business guidelines based on the situation. Application of exclusive business guidelines and mapping metadata support in tailoring every scenario [9]. Delivering a precise and adaptable mechanism using different location-based policy implementations is effective.

Table 2: Data formatting

Module	Functionality	Efficiency and Impact
Unified Medicaid Extract Core	Harmonizes ICN, ACHN, PERM, MMIS, Approved Pay, and other datasets across states	Eliminates redundant engineering and manual formatting efforts
AI Schema Translator	Learns state-specific formats and maps to federal standards dynamically	Accelerates onboarding and reduces mapping errors
Reusable ETL Framework	Modular pipelines for daily, monthly, quarterly exchanges	Saves millions in development hours and ensures audit-readiness
Compliance & Audit Validator	Auto-validates extracts against CMS, PERM, and MMIS rules	Prevents audit penalties and ensures regulatory alignment
Executive Oversight Dashboard	Tracks extract status, error rates, and reuse metrics	Enables strategic funding decisions and inter-agency transparency
Taxpayer ROI Tracker	Quantifies cost avoidance and reuse benefits	Demonstrates public value and supports grant justification
Multi-State Deployment Engine	Rapid onboarding of new states with minimal rework	Promotes shared infrastructure and federal-state collaboration
Secure Data Exchange Layer	FIPS-compliant, encrypted transfers with role-based access	Ensures HIPAA compliance and federal security standards

2.3. Automating regulations

AI-motivated validation allows for enriching to increase the capabilities of content prior to extraction effectively. Automating the process and detecting anomalies is practical for checking randomness and process flows. [10] This allows

decreasing the involvement of errors and enhancing dependability. Delivering maximum precision is effective for enhancing technology adaptability and achieving better outcomes.

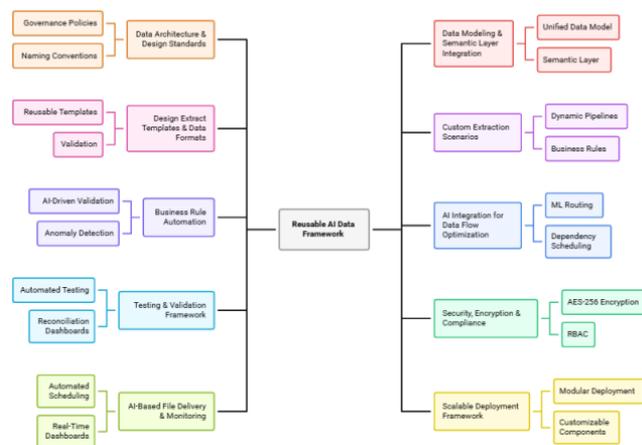


Fig 1: Framework

2.4. Following compliance guidelines

The addition of compliance rules in the automated system allows direct identification and capture needed data from documents available. This ensures comprehensive reporting in accordance with the requirements. Ascertaining this level of accuracy is effective for managing fiscal scenarios and executing required solutions for flagging challenges. These include complying with HIPAA guidelines and obtaining benefits with maximum security [11]. AI data processes are beneficial for capturing data automatically for complete auditing and documentation of details. These allow continuous analysis and public data security.

3. Hybrid approach

There are many advanced tools available to increase the efficiency of AI policy and promote Medicaid services. This includes implementing a combination of AI and human analysis mechanisms. This looping allows increased accuracy and continuously increasing system productivity. The processes enable achieving high levels of accuracy through complete automation [12]. This increases validation capabilities and accuracy with ongoing learning scenarios. Increasing the accuracy of support systems to refine future analytics by correcting processes. These allow managing processes with high-level user oversight for quality and continuous interactions.

4. Successful adaptation of a reusable AI framework for State Medicaid department transactions

Implementing AI-oriented content extraction effectively requires comprehensive preparation across diverse Medicaid data processing dimensions. These processes are initiated by assessing the system preparation levels and the existing document processing flow. These allow setting up volume-based measures and noticing challenges along with possible improvement paradigms. The process flows are effective because they involve different teams and follow technology guidelines. The processes impact business units distinctively [13]. These require high-level sample document creation for training the systems and conducting tests. The table depicts the application of reusable AI for State Medicaid departments.

The mechanism is effective for creating ramny across the preparation activities, which include assessing unification mechanisms and defining process size requirements. The selection of suitable deployment methods is important for achieving maximum impact from these processes.

Table 1: Medicaid applications of Reusable AI

Dataset	Primary Departments	Usage Purpose	Unified Extract Benefit
ICN (Internal Control Number)	State Medicaid, CMS, OIG	Tracks individual claims for audits, fraud detection, and payment reconciliation	Ensures consistent claim traceability across states and federal audits
ACHN (Accountable Care Health Network)	State Medicaid, CMS Innovation Center	Monitors provider networks, care coordination, and value-based payment models	Harmonizes provider identifiers and care metrics for cross-state analysis
PERM (Payment Error Rate Measurement)	CMS, State Medicaid QA Teams	Measures improper payments and compliance with federal rules	Automates extract validation to reduce PERM audit failures
MMIS (Medicaid Management Information System)	State Medicaid IT, CMS Systems Group	Core system for claims processing, eligibility, and provider enrollment	Enables modular reuse of MMIS extracts for federal reporting and analytics
Approved Pay	State Finance, CMS Financial Management	Tracks authorized payments to providers and managed care organizations	Standardizes payment extracts for budget forecasting and federal match validation

5. Extraction stages

Establishment of synergy and creating high-level process flows by minimizing user involvement and errors. The addition of semantics allows collaboration between invoice information and department details, with continuous automation and validation against claims processed and accounting [14]. These continue to promote approvals and initiate fund releases. The reports developed could be aligned with standards established suitably.

5.1. Designing and Extraction

The state-wise and federal policy development initiatives for exploration of opportunities for enhancing robustness, as well as implementing Medicaid management

using AI. The research insights depicted regarding stakeholders' effective use of information management standards. These empower adopting technical advancements responsibly. Increased federal technology support for Medicaid divisions and supporting state-wise learning scenarios. These finally enable efficient use of AI in administering Medicaid processes [15]. Using AI responsibly and increasing success requires the following recommendations

5.2. Continuous review and transparency

AI logic and programming support Medicaid processing efficiency. This could be further enhanced by continuously reviewing and setting up plans for refinement. Transparency enhances trust and indicates that processes are transparent.

Successfully executing these Federal healthcare services and supporting programs requires analysis in multiple dimensions [16]. AI implementation requires appropriate training and transparency to ensure successful execution in these federal-level health support programs. The mechanisms enhance statutory capability to address the distinctive requirements of users in the involved processes.

5.3. Testing for launching processes

Limited range testing processes involved with AI use in Medicaid activities before launching directly. These facilities could support Medicaid-empowered users for developing needed tools for conducting usability and access procedures. The review processes may support checking applications across various types of users. This could be followed by developing a complete app for the launch process [17]. This increases confidence, as advanced systems work in accordance with accessibility guidelines and reduce challenges associated with indirect biases.

5.4. Securing the most vulnerable information

Following these stages supports effective health insurance operations, supports citizens according to eligibility, and assesses and observes AI and algorithm-based systems to ensure the absence of reinforcing discriminatory trends. These commonly eliminate individuals with issues [17]. Automated processes with data-reuse code support these by reducing process length and increasing access. These require avoiding hindrances caused by these scenarios.

6. Strategic influence

Intelligent content processing and AI reuse provide many opportunities to optimize processing activities in Medicaid systems. This allows for reducing the time required for processing and manual data entry, thereby decreasing operational lag. The activities predominantly affect data quality across platforms, including healthcare centers across states and insurance service providers.

Table 4: Strategic Influence

Benefit	Description
Taxpayer Savings	Millions saved annually through reusable templates and automation
Efficiency Gains	Reporting cycles reduced from weeks to days
Compliance Assurance	HIPAA / CMS / state standards embedded at every stage
Scalability	Framework supports expansion across 50 states without reimplemention

7. Recommendations

7.1. Collaborating extraction and automation

AI-motivated connect extraction activities indicate tasks beyond simple ecology updates. These support as a basis for effective Medicaid claim processing and generating maximum value. Leaders use this mechanism to extend document process flows and access content [3]. The high-level capabilities of AI content extraction and intelligent Medicaid claim processing could be realized by adopting robust process flows. These reduce errors and capture the needed content efficiently.

7.2. Implementing data stewardship

This is effective for proactively implementing responsible content re-utilization policies to catalyze and combine efforts to address the challenges involved, and for embedding re-utilization processes into pursuits. This allows for equalizing demand and data accessibility for public processes through continuous iteration. These processes could be embedded to enhance public value and manage content accessibility expertise. The opportunities involved could be identified for combining cross-functional components and allow function-based accessibility.

7.3. Depicting the value proposition of content

This process includes verifying and eliciting stakeholder perceptions of the value generated by data, which is another recommended method. This increases associate participation in data reuse processes as shared across regular departments collaboratively [19]. This , fostering suitable associations and a combination to create the required business scenarios. These supports building a strong case for social benefits.

5.4. Technology components

Implementing AI and a supervised learning mechanism enables the detection of randomness in content. This follows the development of predictive algorithms for compatible forecasts, the automation of business regulations, and continued analytics. Python could be used for secure, capable data extraction, followed by processing and loading into systems. These are automated effectively using Python code. Creating data validation code and streamlining automatic process flow. Using Azure enables orchestrating cloud-native processes and increasing compliance with guidelines.

6. Conclusion

AI could be regarded as the most empowering tool for effectively managing issues occurring in implementing Medicaid policies. Taxpayers could benefit from this app by preventing fraud and redirecting resources to address genuine needs. The upcoming AI support for Medicaid relies on learning algorithms from various dimensions. These federal-level policy developers continuously innovate and are associated with continuous oversight, and clear data management empowers those responsible for policy execution. AI-Reuse frameworks enable delivering program benefits with maximized accessibility and promote affordable healthcare for users with the highest service requirements.

References

1. S. R. Abbas, Z. Abbas, A. Zahir and S. W. Lee, "Federated Learning in Smart Healthcare: A Comprehensive Review on Privacy, Security, and

- Predictive Analytics with IoT Integration," *Healthcare*, vol. 12, no. 24, pp. 1-33, 2025.
2. S. Anand and S. Sharma, "Self-Healing Data Pipelines for Handling Anomalies in Medicaid and CHIP Data Processing," *International Journal of AI, Big Data, Computational and Management Studies*, vol. 5, no. 2, pp. 27-37, 2024.
3. Carbonaro, L. Giorgetti, L. Ridolfi, R. Pasolini, A. Pagliarani, M. Cavallucci, A. Andalò, L. D. Gaudio, P. D. Angelis, R. Vespignani and N. Gentili, "From raw data to research-ready: A FHIR-based transformation pipeline in a real-world oncology setting," *Computers in Biology and Medicine*, vol. 197 B, no. 1, pp. 1-14, 2025.
4. S. Dahl, M. Bøgsted, T. Sag and C. Vesteghem, "Performance of Natural Language Processing for Information Extraction From Electronic Health Records Within Cancer: Systematic Review," *JMIR Publications*, vol. 13, no. 1, pp. 1-13, 2024.
5. R. Eden, I. Chukwudi, C. Bain, S. Barbieri, L. Callaway, S. d. Jersey, Y. George, A.-D. Gorse, M. Lawley, P. Marendy, S. M. McPhail, A. Nguyen, M. Samadbeik and C. Sullivan, "A scoping review of the governance of federated learning in healthcare," *npj Digital Medicine* volume, vol. 8, no. 427, pp. 1-10, 2025.
6. S. Essaid, J. Andre, I. M. Brooks, K. H. Hohman, M. Hull, S. L. Jackson, M. G. Kahn, E. M. Kraus, N. Mandadi, A. K. Martinez, J. Y. Mui, B. Zambarano and A. Soares, "MENDS-on-FHIR: leveraging the OMOP common data model and FHIR standards for national chronic disease surveillance," *JAMIA Open*, vol. 7, no. 2, pp. 1-12, 2024.
7. M. Fruchart, P. Quindroit, C. Jacquemont, J.-B. Beuscart, M. Calafiore and A. Lamer, "Transforming Primary Care Data Into the Observational Medical Outcomes Partnership Common Data Model: Development and Usability Study," *JMIR Med Inform*, vol. 12, no. 1, pp. 1-12, 202.
8. K. Khan, R. Wang, M. Skanderson, C. A. Brandt, S. Fodeh and J. A. Womack, "A Roadmap to Artificial Intelligence (AI): Methods for Designing and Building AI ready Data to promote Fairness," *Journal of Biomedical Informatics*, vol. 154, no. 1, pp. 1-9, 2024.
9. P. Leibo Liu, M. Victoria Blake, B. Matthew Barman, P. Blanca Gallego, M. Timothy Churches, P. Georgina Kennedy, M. Sze-Yuan Ooi, M. P. Geoffrey P Delaney and P. Louisa Jorm, "Using natural language processing to extract information from clinical text in electronic medical records for populating clinical registries: a systematic review," *Journal of the American Medical Informatics Association*, vol. 176, no. 1, pp. 1-16, 2025.
10. K. Adams and A. Buffett, "AI and Medicaid: Balancing the Promise of Efficiency with Guardrails to Ensure Responsible Use," *Bipartisan Policy Center*, 22 October 2025. [Online]. Available: <https://bipartisanpolicy.org/article/ai-and-medicaid-balancing-the-promise-of-efficiency-with-guardrails-to-ensure-responsible-use/>. [Accessed 25 November 2025].
11. Pathak, L. Serrer, D. Zapata, R. King, L. B. Mirel, T. Sukalac, A. Srinivasan, P. Baier, M. Bhalla, C. David-Ferdon, S. Luxenberg and A. V. Gundlapalli, "Privacy preserving record linkage for public health action: opportunities and challenges," *J Am Med Inform Assoc*, vol. 31, no. 11, p. 2605–2612, 2024.
12. P. Tabari, G. Costagliola, M. D. Rosa and M. Boeker, "State-of-the-Art Fast Healthcare Interoperability Resources (FHIR)–Based Data Model and Structure Implementations: Systematic Scoping Review," *JMIR Medical Informatics*, vol. 12, no. 1, pp. 1-27, 2024.
13. H.-A. Talvik, M. Oja, S. Tamm, K. Mooses, D. Särg, M. Lõo, Ö. R. Siimon, H. Šuvalov, R. Kolde, J. Vilo, S. Reisberg and S. Laur, "Repeatable process for extracting health data from HL7 CDA documents," *Journal of Biomedical Informatics*, vol. 161, no. 1, pp. 1-11, 2025.
14. L. Teo, L. Jin, S. Li, D. Miao, X. Zhang, W. Y. Ng, T. F. Tan, D. M. Lee, K. J. Chua, J. Heng, Y. Liu, R. S. M. Goh and D. S. W. Ting, "Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture," *Cell Reports Medicine*, vol. 5, no. 2, pp. 1-13, 2024.
15. N. Williams, "Building the observational medical outcomes partnership's T-MSIS Analytic File common data model," *Informatics in Medicine Unlocked*, vol. 39, no. 1, pp. 1-7, 2023.
16. M. Yvonne Wieland-Jorna, M. Daan van Kooten, R. A. Verheij, P. Yvonne de Man, A. L. Francke and P. Mariska G Oosterveld-Vlug, "Natural language processing systems for extracting information from electronic health records about activities of daily living. A systematic review," *JAMIA Open*, vol. 7, no. 2, pp. 1-10, 2024.
17. S. K. Chebrolu, "AI-Powered Business Intelligence: A Systematic Literature Review On The Future Of Decision-Making In Enterprises," *SSRN*, vol. 2025, no. 1, pp. 1-30, 2025.
18. T. Cho and B. J. Miller, "Using artificial intelligence to improve administrative process in Medicaid," *National Center for Biotechnology Information*, vol. 2, no. 2, pp. 1-10, 2024.
19. S. Dandolu, "Cloud-Native Architecture for AI Data Platforms: A Snowflake Implementation Case Study," *World Journal of Advanced Engineering Technology and Sciences*, vol. 15, no. 3, pp. 475-485, 2025.
20. S. Dashora, N. Gupta, M. P. S. Tanwar, A. Singh and S. Kaur, "Cloud-based Data Analytics for Business Intelligence," *International Journal for Research in Applied Science & Engineering Technology*, vol. 11, no. 11, pp. 2758-2765, 2023.
21. F. Dakalbab, M. A. Talib, Q. Nasir and T. Saroufil, "Artificial intelligence techniques in financial trading: A systematic literature review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 3, pp. 1-10, 2024.