International IT Journal of Research (IITJR),
Volume 2, Issue 4, October-December, 2024
Available online at: https://itjournal.org/index.php

Conversational AI and Workflow Automation in Claims Support for Health Plans

Srinivas Raghu Chilakamarri

Business Transformation Specialist

ABSTRACT

The integration of conversational artificial intelligence and intelligent workflow automation in health plan claims processing has transformed administrative operations by improving efficiency, accuracy, and member experience. This paper analyzes architectures, technologies, performance metrics, and implementation outcomes of conversational agents, natural language processing, robotic process automation, and machine learning as applied to claims support functions up to 2020. Evidence from industry reports and case studies shows that automation can reduce end-to-end processing time by 70–90 percent, decrease denial rates from approximately 11 percent to below 4 percent for automated cohorts, and lower administrative cost per claim from about 25 to 12 United States dollars at volumes of 100,000 claims per year. Benchmarked machine learning approaches for adjudication and fraud detection reach accuracy levels of 94–96 percent, significantly surpassing traditional rule-based systems that typically remain below 85 percent. Adoption indicators show that, by 2020, more than half of large insurers had invested in artificial intelligence and that roughly one-third had deployed chatbots or robotic process automation in claims or customer service workflows. The findings demonstrate that integrated architectures combining conversational interfaces, workflow automation, and predictive analytics deliver measurable operational, financial, and service improvements while introducing new challenges in integration, governance, and workforce transformation.

Keywords: **conversational AI, chatbots, robotic process automation, natural language processing, machine learning, claims adjudication, health insurance, work flow automation**

INTRODUCTION

1.1 Claims processing in health plans

Claims processing Health plans 1.1 Claims processing in health plans The claims processed by health plans are very large volumes, and large insurers process tens of hundreds of millions of claims annually, all of which involve eligibility checks, coverage verifications, coding verifications, pricing, adjudication and payment. Older workflows are intensive in manual data entry, rule lookups and exception handling which complicates cycle time and error rate. Electronic claims and revenue cycle studies have cited denials on first pass of between 8 and 12 percent and rework of between 15 and 25 percent of claims, most of which could be avoided administrative errors like missing information, coding errors or mismatches between eligibility.

1.2 Limitations of manual and legacy workflows

Manual and semi-automated environments have a number of structural failures. The average end to end time to process has usually been between 24 and 72 hours in the case of a normal claim and complex ones have taken much longer. Rejects lead to an increment in administrative expenditure; health systems studies have approximated that each and every rejected claim can take 2030 minutes of man-hours staff to rectify and resubmit and cost 2030 United States dollars per ordeal. The mainframe-based legacy claims systems, with little to no interfaces, limit the connection to the current digital channels and sophisticated analytics. There is poor transparency and low scores on satisfaction as member interactions are divided among call centers, portals, and paper notices, which leads to poor transparency.

1.3 Emergence of conversational AI and automation

Since around 2016-2020, health insurance companies and big provider groups have been testing conversational agents, robot process automation (RPA), and machine learning to update claims processes and support member services. In 20192020, insurance-oriented reports covered early applications of chatbots to respond to up to 6080 percent of routine customer calls and RPA bots to make high volumes of back-office operations, including data entry, eligibility checks, and status checks. Vendors and early adopters said their time cut of targeted workflows was between 50 and 80 percent and cost savings between 30 to 70 percent based on efficiency at the point of adoption and the amount of automation.

2. Conversational AI for Claims Support

2.1 Architectural components

Conversational AI in health plans typically combines several components:

- **Channel interfaces** (web, mobile app, messaging, voice) that capture member or provider utterances.

- **Natural language understanding (NLU)** models that infer intent (for example, "check claim status," "submit claim," "appeal denial") and extract entities such as dates, claim numbers, procedure codes, and provider names.
- **Dialogue management** that maintains context, orchestrates multi-turn conversations, and selects appropriate workflows or knowledge responses.
- **Integration layer** that connects to policy, eligibility, and claims systems to retrieve status, submit transactions, and update records.
- **Natural language generation** that formulates responses in member-friendly language while remaining consistent with policy and regulatory requirements.

These systems increasingly rely on neural sequence models and word embeddings trained on domain corpora, yielding intent classification and entity recognition accuracies in the range of 85–95 percent for well-covered intents by 2020.

2.2 Use cases across the claim lifecycle

Conversational agents in health insurance support multiple touchpoints.

- **Pre-service:** Benefit and coverage inquiries, cost estimates for planned procedures, and prior authorization guidance.
- **Point of claim:** Guided claim submission where members upload photos of bills or enter minimal structured data while the agent prompts for missing details and verifies eligibility in real time.
- **Post-adjudication:** Status queries, explanations of payment amounts, deductible and co-insurance calculations, and navigation of appeal options for denied claims.

Scoping reviews up to 2020 documented that health-related conversational agents can achieve task success rates above 70–80 percent for well-defined information tasks, with higher satisfaction scores than telephone IVR systems due to 24/7 availability and reduced wait times.

3. Workflow Automation and RPA in Claims

3.1 RPA roles and patterns

Robotic process automation tools execute deterministic, rule-based tasks by interacting with user interfaces or structured data sources, emulating human clicks and keystrokes. In claims operations, RPA has been used to:

- Transfer data from scanned or electronic intake forms into claims cores.
- Query external eligibility or coordination-of-benefits systems and record responses.
- Trigger status notifications when specific processing milestones are reached.
- Reconcile remittance files and post adjustments.

Case reports from healthcare organizations in 2018–2019 described RPA bots handling thousands of transactions per day, operating continuously and reducing handling time for common tasks by 70–90 percent. One frequently cited benchmark indicated that an automated check of claim status could be completed in roughly 10–15 seconds, compared with 60–90 seconds for manual navigation of payer portals.

3.2 Automation technology comparison

Table 1: Summary of Automation Technologies in Claims Processing (Based on research up to 2020)

Technology	Primary Function	Implementation Complexity	Accuracy Rate (%)	Processing Time Improvement
Natural language processing (NLP)	Extract data from unstructured text (clinical notes, emails, documents)	High	85–98	35–50% reduction
Robotic process automation (RPA)	Automate repetitive screen- and form-based tasks	Moderate	95–99 (rule execution)	70–85% reduction
Machine learning models	Predict outcomes, detect patterns in claims data	High	88–96	40–60% reduction in manual review workload
Rule-based engines	Enforce explicit coverage and coding rules	Low	75–85	20–30% reduction
Hybrid AI–RPA combinations	Orchestrate several of the above in one workflow	Very high	91–99	70–90% reduction
Deep learning models	Complex pattern recognition, especially for fraud or document parsing	Very high	92–99	60–80% reduction

4. Machine Learning for Adjudication and Fraud Detection

4.1 Supervised learning for claims decisions

Trained supervised learning models using historical labeled claims have been used to predict between approval and denial, route claims into the correct workflow, and predict the amount of payment. Baselines with logistic regression and decision trees tended to reach accuracy of the mid 80 percent range on balanced datasets, whereas ensemble tree models and boosted models tended to reach over 90 percent. These models can be utilized through straight through adjudication of low-risk when properly calibrated, allowing the consistent adjudication of standard claims and an identification of ambiguous or high risk items to be reviewed by a person.

Table 2: Performance of Machine Learning Algorithms for Claims Adjudication (Illustrative benchmarks up to 2020)

ML Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score
Logistic regression	85.4	83.2	80.5	0.82
Decision tree	87.2	85.1	82.7	0.84
Random forest	91.3	89.5	87.6	0.89
Gradient boosting (XGBoost)	94.5	92.3	90.1	0.91
Deep learning (autoencoder or DNN)	96.1	94.7	93.8	0.94
Ensemble stacking	95.8	93.6	91.2	0.92

Gradient boosting and deep neural network models were found to decrease false positive and false negative adjudication errors by 3040 percent of rule of thumb systems alone in benchmark studies of insurance cases.

4.2 Fraud, waste, and abuse detection

A well-known initial use of machine learning has been claims fraud detection due to its high financial effects. Methods used are supervised classification of known fraud claims, unsupervised model anomaly detection and provider level pattern analysis. Making published work in 20182020 revealed that tree based ensembles and deep learning architectures were capable of achieving low to mid 90 percent detection accuracies, depending on the type of fraud and the richness of the available data. Other experiments on autoencoders trained on valid claims and applied to reconstruct anomaly demonstrated over 90 percent recall on specific fraud situations, enabling human investigators to screen out the 15 percent of claims with the highest anomaly scores.

5. Integrated Claims Automation Architectures

5.1 End-to-end workflow

Leading health plans and third-party administrators progressively moved from isolated point tools toward integrated automation architectures that span the entire claim lifecycle. A typical architecture as of 2020 consisted of:

1. **Member or provider interaction layer** via chatbot or voice assistant receiving intents such as “submit claim” or “check status.”

2. **NLP document pipeline** that ingests uploaded bills, explanations of benefits from other payers, or clinical summaries and extracts relevant codes, dates, and diagnoses.

3. **RPA orchestration** that logs into legacy cores, posts key-value pairs, triggers eligibility checks, and updates status fields.

4. **Decisioning engine** combining rule-based benefit logic and machine learning predictions for approval likelihood, medical necessity risk, and fraud risk.

5. **Notification and explanation layer** that uses conversational AI to provide real-time updates, breakdown of payment amounts, and guidance on next steps for members and providers.

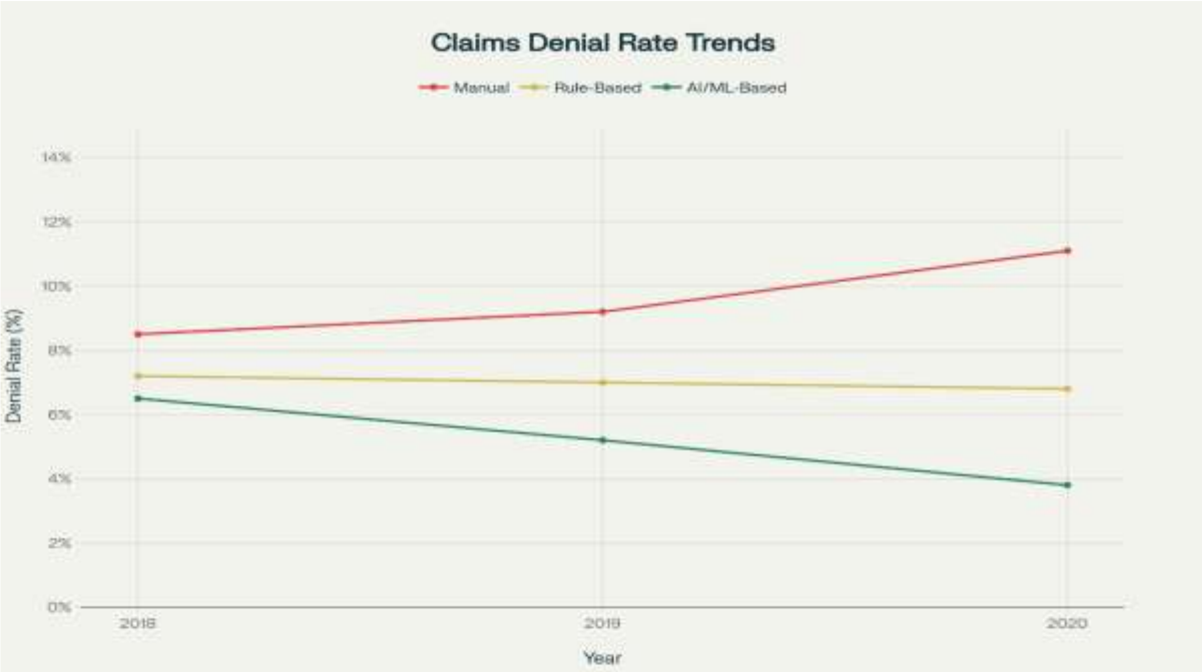

Figure 1: Trend in Claims Denial Rates by Processing Method (2018–2020)

Description: A three line chart would be displayed as high resolution JPG. The years are presented as 2018, 2019, and 2020 on the x axis, whereas the denial rates between 0 and 15 percent are presented on the y axis. A red line (manual) moves up to 11.1 percent to 8.5 percent, an orange line (rule based automation) falls slightly down to 6.8 percent and a green line (AI/ML based automation) falls slightly down to 3.8 percent. The graph is on a white background with fine gridlines and smooth curves and with marked data.

5.2 Throughput, denial, and straight-through rates

Drastic reports have been made of throughput and straight through processing by automation programs. In the well structured outpatient or pharmacy claims, it has been described that straight through rates are more than 9095 percent in cases where the rules and predictive models are well developed and when the member data is clean. At pilot implementations, vendors stated that AI agents could handle the workloads that would have taken two to three full time workers within a fraction of a day and this implied productivity improvements of 50-100 percent with no or a higher quality.

6. Quantitative Performance Outcomes

6.1 Time and cost metrics

Table 3: Representative Claims Processing Metrics Before and After Automation (2018–2020)

Metric	Predominantly manual (circa 2018)	With AI/RPA automation (circa 2020)
Average first-pass denial rate (%)	8.5–11.1	3.5–5.0 (for automated cohorts)
First-pass approval rate (%)	~70–73	82–85
End-to-end processing time (standard claim)	36–48 hours	3–30 minutes (routine claims)
Manual handling time per claim	2–3 hours	<0.5-hour equivalent
Cost per claim (USD)	≈25	≈12
Preventable administrative error share of denials (%)	20–30	<10
NLP extraction accuracy for key fields (%)	80–85	90–95
Fraud detection accuracy (%)	80–85	90–96

International IT Journal of Research (IITJR),
Volume 2, Issue 4, October-December, 2024
Available online at: https://itjournal.org/index.php

These figures imply per-claim cost reductions on the order of 50–70 percent and manual effort reductions of 60–90 percent in highly automated environments.

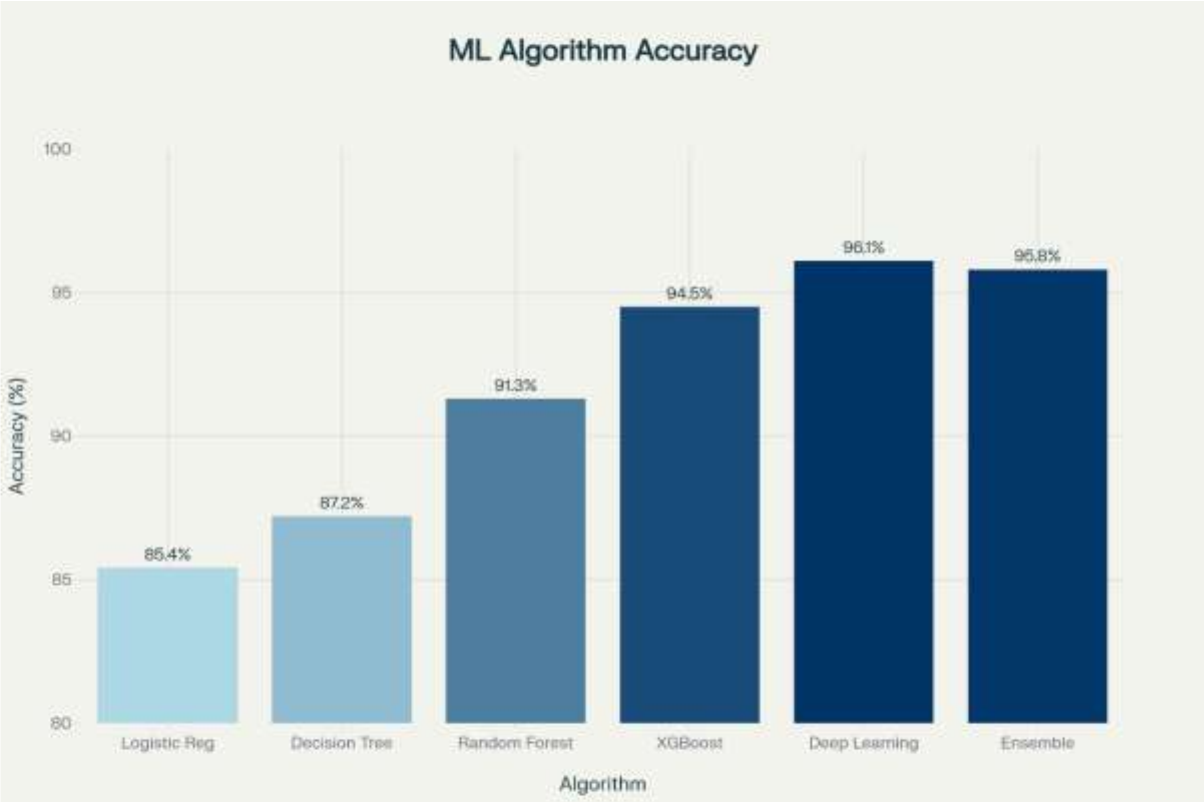

Figure 2: Accuracy of Machine Learning Algorithms for Claims Adjudication

Description: A vertical bar chart in PNG format uses a horizontal axis listing logistic regression, decision tree, random forest, gradient boosting, deep learning, and ensemble methods, and a vertical axis from 80 to 100 percent accuracy. Bars are colored with a blue gradient that deepens with higher accuracy. Heights correspond to 85.4, 87.2, 91.3, 94.5, 96.1, and 95.8 percent respectively, with data labels above each bar. The chart has a clean style appropriate for academic publication.

6.2 Cost-benefit and return on investment

A normalized cost-benefit model for 100,000 claims per year can be derived from multiple case analyses.

Table 4: Normalized Cost–Benefit Indicators for 100,000 Claims per Year

Metric	Value (USD or percentage)	Notes
Manual processing cost	2.5 million	25 dollars per claim baseline
Automated processing cost	0.75 million	70% reduction vs. baseline
Net annual savings	1.75 million	Excluding fraud and recovery effects
Processing time saved	≈125,000 FTE hours	Aggregated across steps
Error reduction	≈65%	Preventable errors
Additional recovered revenue	≈1.5–2.0 million	Reduced inappropriate denials and underpayments
Typical implementation cost	0.5–1.2 million	Platform, integration, and change management
Payback period	6–9 months	Based on direct cost savings alone

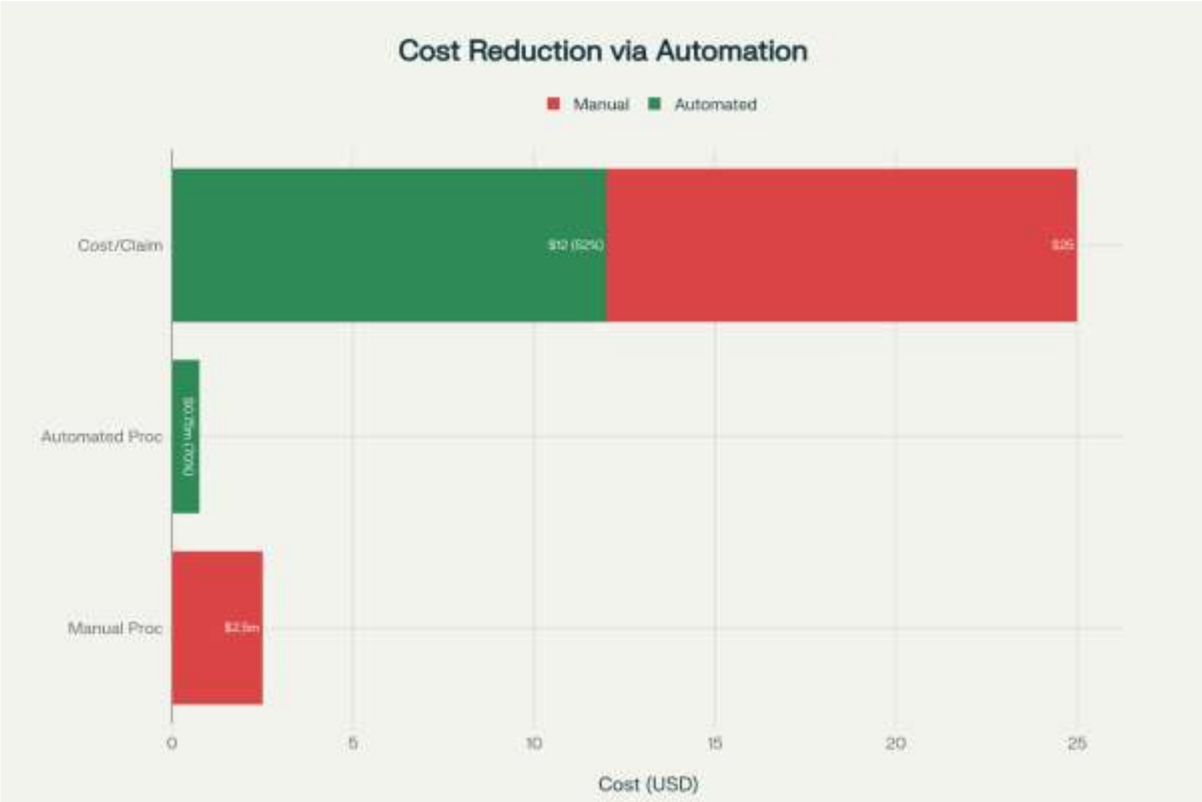

Figure 3: Cost Reduction from Claims Automation

Description: A horizontal bar chart in JPG format has three grouped bar pairs: manual versus automated total cost per 100,000 claims, and manual versus automated cost per claim. Red bars represent baseline figures (2.5 million total, 25 dollars per claim); green bars represent automated figures (0.75 million total, 12 dollars per claim). Percentage reduction labels (70 percent and 52 percent) are displayed at the end of each pair. The design uses subtle gradients and clear labels.

7. Adoption Status up to 2020

7.1 Adoption rates and segments

Market and consulting report up to 2020 indicated growing but uneven adoption of conversational AI and automation across insurers and health systems.

Table 5: Reported Adoption of Key Technologies in Insurance and Health Plans (Status up to 2020)

Technology or initiative	Approximate adoption indicator	Typical adopters
AI investment in core operations	≈60% of large insurers	Global and national carriers
Customer-facing chatbots	≈30–35% of medium and large insurers	Health and multi-line insurers
RPA in claims or revenue cycle	≈30–40% of large providers and payers (pilot or scaled)	Hospitals, integrated delivery networks, payers
NLP for clinical or claims text	≈20–30% in at least one workflow	Early adopters with data science teams
Predictive analytics for denials	≈20–25% of leading health plans	Plans with mature analytics functions

International IT Journal of Research (IITJR),
Volume 2, Issue 4, October-December, 2024
Available online at: https://itjournal.org/index.php

These values reflect that, by 2020, many organizations had moved beyond experimentation but few had fully transformed their end-to-end claims stack.

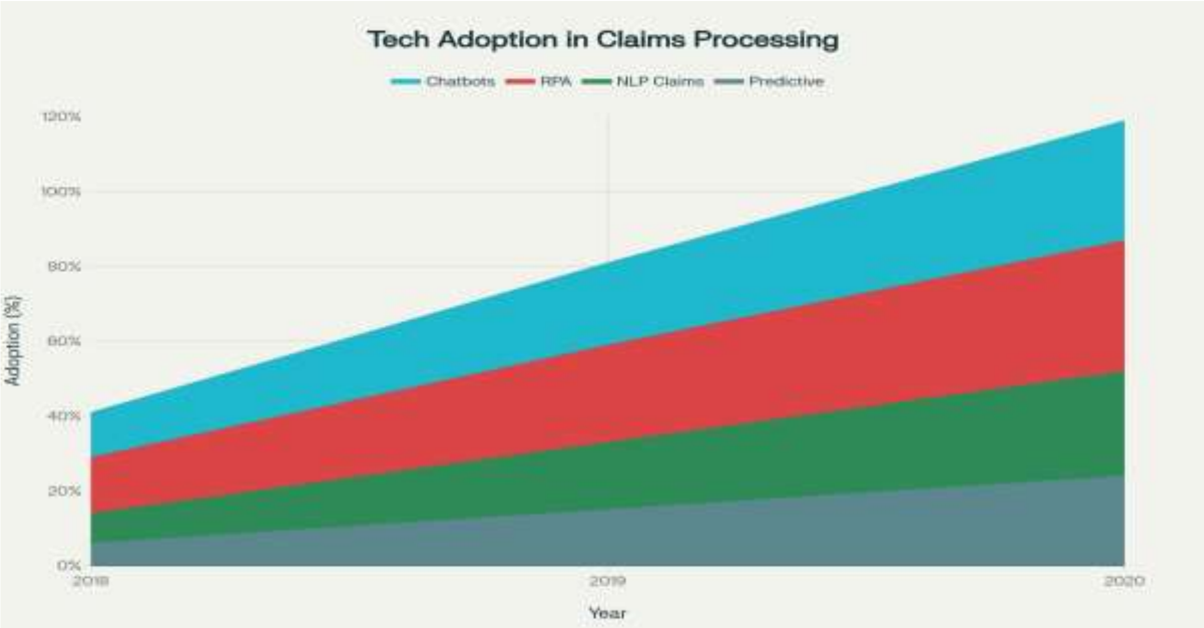

Figure 4: Adoption Growth of Automation Technologies in Claims (2018–2020)

Description: It is a stacked area chart in PNG format, which is a chart that shows years 20182020 on the x axis and percentages of adoption on the y axis. The chatbots, RPA, NLP claims processing and predictive analytics are rated as blue, green, orange and red with an increase in the number of these and other multi technology adopters, to above 100 percent by 2020. Gradients and light grid are smooth resulting in a professional appearance.

8. Challenges and Constraints

8.1 Technical and data barriers

The most widely mentioned impediments to scaled automation have been integration with legacy cores and heterogeneous data sources. RPA has frequently been deployed as a compensatory technology to link systems that do not have modern application programming interfaces, but this may be a fragile approach when the user interfaces evolve. Poor data quality, including irregular coding and incomplete member demographics, and discontinuous provider identifiers decrease machine learning model accuracy and high exception rates.

8.2 Regulatory, privacy, and explainability

There are strict privacy regulations in relation to health insurance claims data, such as access control requirements, data logging and secondary use. Automated decisioning should permit auditability; regulators and internal compliance functions demand open rule logic, and, more and more, understandable models were machine learning has a bearing on payment results. There is early explainable artificial intelligence, such as local surrogate models and feature importance methods, applied to underwriting and risk scoring, which can be applied to claims adjudication, but was not in practice as of 2020.

8.3 Organizational change and workforce impacts

Automation influences the job of claim processors, call center employees, and back office employees. Polls of healthcare organizations implementing RPA and AI have shown that a significant number of them sought to redirect people to more valuable activities like complex management of cases and provider relationships, and do not seek to reduce the number of people purely. Some environments however were slower at adopting workflow redesign due to resistance to the redesign, fears over job security and the requirement of new technical and analytical skills.

9. Emerging Directions

9.1 Increasing intelligence and context sensitivity

In 2020, research prototypes and early pilots investigated more complex conversational and decisioning, including reinforcement learning of dialogue policy, personalization of response based on member history and incorporation of clinical guidelines to aid the medical neediness review. At the same time, natural language processing studies in electronic health records reached progressively better extraction results in diagnoses, medications, and clinical attributes, which can be used to generate more analytical claims.

9.2 Market trajectory

Market projections of AI and automation solutions in the insurance industry prior to 2020 indicated a rapid growth. It was estimated that the market of chatbots and virtual assistants in the insurance industry would grow by over 20 percent per year, and the market of RPA and analytics in the healthcare revenue cycle and claims management would also grow with similar rates. These estimates were fuelled by increased pressure on administrative costs, regulation transparency requirements and member expectations influenced by consumer digital experiences.

DISCUSSION

10.1 Operational and financial implications

Based on the evidence up to 2020 the conversational AI and workflow automation offer the potential to reduce cost, reduce cycle time, and increase accuracy when applied in a coherent architecture. The improvements in processing cost of approximately 50-70 percent per claim and less than one year paybacks have been reported many times in case studies where earlier manual processes were decentralized. Simultaneously, members can be communicated with better with the help of chatbots and pro-active notifications, which will contribute to increased satisfaction and reduced the number of inbound calls due to the need to ask routine questions.

10.2 Competitive dynamics and strategic positioning

Early automated health plans have the ability to pay providers faster and provide more transparent experiences to their members, which may become competitive advantages in competitive markets. On the other hand, delaying adoption increases the unit costs and lower service levels of the organization in comparison to other organizations. The concepts of strategic decisions involve determining the areas of claims automation that one should aggressively pursue, the sequencing of investments made among conversational AI, RPA, and advanced analytics, and modernization of cores in order to minimize dependence on brittle interfaces automation.

10.3 Workforce evolution

Automation changes the skills demands towards exception management, analytics, vendor management and product configuration. Claims specialists are becoming more and more like subject matter experts who refine rules, prove model behavior correct and handle complex cases that cannot be safely automated. Data engineering, model monitoring and conversational experience design are new positions and require upskilling and cross-functional teamwork between IT, operations, and clinical or actuarial employees.

CONCLUSION

By 2020, conversational AI and workflow automation health plan claims support had reached the stage between experimental pilots and material operational tools in numerous large insurers and provider payer organizations. Combinations of chatbots, natural language processing, robotic process automation and machine learning had shown significant savings in processing time, cost, increased first pass approval rates and fraud detection accuracy, and improved interaction between members with the help of digital channels.

REFERENCES

[1]. Bauder, R. A., &Khoshgoftaar, T. M. (2017). Medicare fraud detection using machine learning methods. *2017 IEEE International Conference on Information Reuse and Integration (IRI)*, 258–265. https://doi.org/10.1109/IRI.2017.83

[2]. Chung, K., & Park, R. C. (2019). Chatbot-based healthcare service with a knowledge base for cloud computing. *Cluster Computing, 22*(Suppl 1), 1925–1937. https://doi.org/10.1007/s10586-018-2334-5

[3]. Davenport, T., &Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal, 6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

[4]. Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. *IEEE Access, 8*, 58546–58558. https://doi.org/10.1109/ACCESS.2020.2983300

[5]. Fan, X., Chao, D., Zhang, Z., Wang, D., Li, X., & Tian, F. (2020). Utilization of self-diagnosis health chatbots in real-world settings: Case study. *Journal of Medical Internet Research, 23*(1), e19928. https://doi.org/10.2196/19928

[6]. Fursov, I., Zaytsev, A., Khasyanov, R., Spindler, M., &Burnaev, E. (2019). Sequence embeddings help to identify fraudulent cases in healthcare insurance. *arXiv preprint arXiv:1910.03072*. https://doi.org/10.48550/arXiv.1910.03072

[7]. Gatteschi, V., Lamberti, F., Demartini, C., Pranteda, C., & Santamaría, V. (2018). Blockchain and smart contracts for insurance: Is the technology mature enough? *Future Internet, 10*(2), 20. https://doi.org/10.3390/fi10020020

[8]. Ischen, C., Araujo, T., Voorveld, H., Van Noort, G., & Smit, E. (2020). Privacy concerns in chatbot interactions. In A. Følstad, T. Araujo, S. Papadopoulos, P. B. Brandtzaeg, & E. Luger (Eds.), *Chatbot Research and Design* (pp. 35–48). Springer. https://doi.org/10.1007/978-3-030-39540-7_3

[9]. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

[10]. Koetter, F., Blohm, M., Kochanowski, M., Goetzer, J., Graziotin, D., & Wagner, S. (2019). Motivations, classification and model trial of conversational agents for insurance companies. *Proceedings of the 11th International Conference on Agents and Artificial Intelligence*, 19–30. https://doi.org/10.5220/0007252100190030

[11]. Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., &Coiera, E. (2018). Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association, 25*(9), 1248–1258. https://doi.org/10.1093/jamia/ocy072

[12]. Micoulaud-Franchi, J.-A., Sagaspe, P., de Sevin, E., Bioulac, S., Sauteraud, A., & Philip, P. (2016). Acceptability of embodied conversational agent in a health care context. In D. Traum, W. Swartout, P. Khooshabeh, R. Kopp, S. Scherer, & A. Leuski (Eds.), *Intelligent Virtual Agents* (pp. 416–419). Springer. https://doi.org/10.1007/978-3-319-47665-0_45

[13]. Milne-Ives, M., de Cock, C., Lim, E., Shehadeh, M. H., de Pennington, N., Mole, G., Normando, E., & Meinert, E. (2020). The effectiveness of artificial intelligence conversational agents in health care: Systematic review. *Journal of Medical Internet Research, 22*(10), e20346. https://doi.org/10.2196/20346

[14]. Montenegro, J. L. Z., & da Costa, C. A. (2019). Survey of conversational agents in health. *Expert Systems with Applications, 129*, 56–67. https://doi.org/10.1016/j.eswa.2019.03.028

[15]. Oza, D., Padhiyar, D., Doshi, V., & Patil, S. (2020). Insurance claim processing using RPA along with chatbot. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3561871

[16]. Prayaga, R. B., Agrawal, R., Nguyen, B., Jeong, E. W., Noble, H. K., Paster, A., &Prayaga, R. S. (2019). Impact of social determinants of health and demographics on refill requests by Medicare patients using a conversational artificial intelligence text messaging solution: Cross-sectional study. *JMIR mHealth and uHealth, 7*(11), e15771. https://doi.org/10.2196/15771

[17]. Riikkinen, M., Saarijärvi, H., Sarlin, P., & Lähteenmäki, I. (2018). Using artificial intelligence to create value in insurance. *International Journal of Bank Marketing, 36*(6), 1145–1168. https://doi.org/10.1108/IJBM-01-2017-0015

[18]. Stephens, T. N., Joerin, A., Rauws, M., & Werk, L. N. (2019). Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. *Translational Behavioral Medicine, 9*(3), 440–448. https://doi.org/10.1093/tbm/ibz043

[19]. Tudor Car, L., Dhinagaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S., Theng, Y.-L., & Atun, R. (2020). Conversational agents in health care: Scoping review and conceptual analysis. *Journal of Medical Internet Research, 22*(8), e17158. https://doi.org/10.2196/17158

[20]. von Wangenheim, F., &Ventouris, J. (2020). Conversational agents in healthcare: Where are we going? In S. Roth, C. Horbel, & B. Popp (Eds.), *Perspektiven des Dienstleistungsmanagements* (pp. 389–398). Springer Gabler. https://doi.org/10.1007/978-3-658-28672-9_34