ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Advancing Reinsurance with AI-Driven Data Integration and Compliance

Venkata Raja Anil Kumar Suddala

Sr Devops Engineer, Sigma IT Corp, USA

DOI: https://doi.org/10.51583/IJLTEMAS.2026.15020000141

Received: 05 March 2026; Accepted: 10 March 2026; Published: 26 March 2026

ABSTRACT

This study has developed a new way to structure the data used in global life/health reinsurance by converting large cedant bordereaux into a scalable, AI-enabled product that will overcome some of the limitations of traditional DB2 (batch ETL) methods. This new Azure Synapse-focused model complies with both Solvency II/IFRS 17, enables schema-agnostic ingestion processes, allows for the application of Apache Spark transformations, and utilizes Python to process reconciliations. As a result, the analysis of this product shows significant outcomes; a reduction in treaty liability calculation time by 75%, a decrease in reconciliation resources used by 70%, an 80% reduction in the costs associated with preparing for audits, and a 400% increase in ingestion capacity. The innovations originally focused on resolving inconsistencies in the cedant formats, eliminated many of the manual processes, and addressed some of the regulatory barriers, utilizing AI-enabled quarantine queues, real-time catastrophe accumulation processes, and proactive fraud alerts. The evaluation metrics show very high success with an AUC of>0.92 (for real-time fraud detection), data quality metrics exceeding 97%, and an average uptime of 99.99%, thus preparing the platform for future evolutions around GenAI, federated learning, and the development of digital twin risk modeling that can adapt to a rapidly changing marketplace.

Keywords: Cedant Bordereaux, Azure Synapse, Apache Spark Transformations, GenAI

INTRODUCTION

When it comes to reinsurance, assure that you use its safety net effect, allowing insurance companies to share their risk with other reinsurance organizations. This allows large or unpredictable risks (like pandemics or natural disasters) to be transferred to others rather than an insurer alone risking their entire business on one catastrophic event. By sharing this risk across many reinsurers, reinsurance improves a company's overall financial stability and ability to withstand market impacts during disruptive events. Benefits of reinsurance include: the ability to better diversify risk to reduce or eliminate the impact of a catastrophic claim on one insurer; a "safety net"; to allow insurers to hold onto more capital reserves to avoid becoming insolvent; enhancing the capacity of insurers to underwrite large policies; and providing more stability for the entire insurance market by lessening variability in claims. Reinsurance has helped to promote growth, revenue, and retention of clients for smaller insurers, as it has enabled them to effectively compete in today's marketplace [1].

The placement of reinsurance is done through a series of steps that begin with the assessment of potential exposures and determination of the need for reinsurance. Once the initial analysis is completed, the cedent organization will create a proposal or renewal package that includes information about the program, terms, objectives, and risk data that is sent to potential reinsurers for review. The cedent organization will subsequently select which brokers and/or reinsurers to send the proposal to for preliminary quotes. After obtaining the preliminary quotes, the cedent organization will perform an evaluation of these quotes and develop benchmarks against them. Depending on the size of the submission, the primary reinsurer will lead the negotiation phase to finalize and confirm the terms, structure, sharing of risk, and the rates and conditions.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

All required documentation including reinsurance agreements and contracts that outline the formalization of the reinsurance partnership.

A global leader in the business of life and health reinsurance, the Company underwrites many of the significant risk exposures of life and health insurers internationally by processing vast amounts of complex data derived from policies and claims. The business model differs from that of primary insurers in that it provides services only to other businesses (B2B); therefore, it is exposed to a broad range of cedants whose data sets it must validate in order to deliver accurate insurance products. Because even minor discrepancies in the actuarial modeling of an insurer's life or health insurance business could create catastrophic and financially damaging results to a reinsurer, reinsurers must comply with the strict conditions of the insurance regulatory authorities to preserve the integrity of their models and provide the necessary traceability of their data. To compete effectively in the reinsurance market, reinsurers must be able to immediately adapt their reinsurance portfolios and pricing strategies as they see changes in real time due to factors such as natural disasters or the present high unemployment rate. As a result, reinsurers need tools to provide their underwriters with immediate access to the most accurate risk assessment information so that they can adjust their pricing promptly when necessary. For example, tools like Allphins, Swiss Re, etc., are useful for improving speed of data reconciliation during a crisis situation by removing the time-consuming task of manual reconciliation while also providing up-to-date and accurate data for making informed reinsurance pricing decisions [2].

Reinsurance is structured through a complex B2B model. While primary insurers sell retail policies to individual policyholders, reinsurers only deal with other insurers and reinsurers throughout the world who submit bordereaux (BDS) in different formats and quality levels, requiring extensive normalization. Legacy systems and technologies have added to many of these obstacles and created additional delays in processing large volumes of Claims and Policies data, limiting companies' ability to enjoy Fast Reporting and Pricing in the ongoing surge of Global Demand. As the complexity of Regulatory Requirements including Solvency II and IFRS 17 continues to increase Actuaries have to rely heavily on manual reconciliations, which are prone to error. Because of the lack of Transparency in the Audit Process, there is an inherent Increase in the Risk of Mispricing. The requirement for Timely Updates to Real-Time Exposure Data in times of Major Disasters has significantly increased pressure on the Existing Systems. The Batch Processing Delays resulted in losses, as competitors took Advantage of the Failures to Obtain Market Share. Furthermore, irregularity in bordereaux's submission during a time of crisis or a Pandemic results in significant amounts of Manual Data Cleaning, which increases the Risk of Data Loss and increased Regulatory Issues. The Need for Expedite Automation of Compliance Modernization has never been more evident; current Manual Processes produce Non-Scalable Processes and Substantially Increase Error Rates and Regulatory Costs [3].

Based on the increasing Global Demand, the Actuarial Industry had to rapidly Transform due to a multitude of Legacy Bottlenecks, Data Inconsistencies and Regulatory Complexities. Under Solvency II and IFRS 17 regulations, even a Minor Delay or Minor Error could lead to a Massive Financial and Compliance Impact. Major Strategies were to create New Modern Pipelines to provide Real-Time Processing of Cedant Data, automate the Reconciliation of Bordereaux from Multiple Sources, Ensure Compliance through Integrated Lineage Tracking and Audit Trails, and utilize Artificial Intelligence Techniques to Enhance Intelligent Risk Analytics.As the reinsurance market evolves, reinsurers must meet increased pressures to take on roles as risk orchestrators rather than simply providing backup support. As recently as 2024, global uninsured catastrophe losses related to climate-related events were reported at an estimated $250 billion. These changes create new emerging risks that will require the use of AI for predictive modeling in order to anticipate these types of events, as demonstrated by projections of $10 trillion in annual exposures by 2028. There is also the additional risk of increasing political violence in unstable regions resulting in dramatic spikes in claims as well as the emergence of new types of baskets of political/terrorist risks. In order to remain competitive and maintain agility, reinsurers will turn to advanced data architectures in order to create a more efficient way of processing data through the use of AI and machine learning technologies.

Several challenges will confront reinsurers throughout the next several years; specifically, the continued increases to record levels of natural catastrophe losses (over $100 billion annually by 2025), exacerbated by severe weather (e.g., hurricanes, floods, etc.) and social inflation (e.g. accident/torts claims) leading toward

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

self-retention practices for many insurers (captives), which in turn diminishes the relevance of traditional reinsurance. Further compounding this trend is the burgeoning threat of cyber risks that are already showing signs of creating significant additional strain on the existing systems. Reinsurers need to develop what are referred to as hyper-agile infrastructures in order to model risk in real-time, to provide immediate notification of loss, and to increase the likelihood of prevention of any permanent protection gaps within the reinsurance market [4].

Reinsurers are faced with considerable difficulties during peak disaster seasons due to the time lag in real-time pricing resulting from the inherent inefficiencies of legacy systems that can only process massive amounts of historical data via batch processing procedures, resulting in delays between 2-5 days before the pricing of any cat treaty occurs, resulting in their inability to be nimble enough to actively participate in competitive bid processes immediately after such disasters occur. Further complicating the actuarial teams' ability to price treaties are the additional time and resource commitments required to track, maintain, and normalize the historical data that comes from more than 50 different global insurers with no common standardized format, which greatly increases the risk of data loss and essentially guarantees a greater than 25% inflation of error rates for any pricing analysis derived from inconsistent, multiple sources.

During natural catastrophes (e.g. hurricanes) when claiming increases from non-proven claims can jeopardize a reinsurer's solvency ratio, reinsurers will experience additional pressure to process claims in compliance with regulatory requirements and be able to demonstrate the use of traceable data within their systems; otherwise, they may be faced with massive multi-million dollar fines and lengthy audit periods. If reinsurers do not adopt modern, AI-driven infrastructure that can reconcile cat treaty prices in seconds, they will lose valuable competitive agility when the demand for rapid pricing support is the greatest. Additionally, reinsurers will experience increasing protection gaps due to their inability to anticipate the accumulation of risk because they are still utilizing outdated approaches to cat treaty pricing. The industry is continuously evolving, and reinsurers will likely need to invest heavily in enhancing their data architectures to accommodate ongoing volatility in the marketplaces, thereby driving significant insurer movement toward self-insured models for many insured objects [5].

BACKGROUND WORK

to do intelligent discovery and profiling of data to get information efficiently from many disparate file formats and system types, by utilizing various machine learning methods to locate trends and inconsistencies in the different source data (for example, policy systems and claims systems) as they relate to each other. This type of automation improves the speed of converting schema from one system to another, which also greatly reduces or eliminates the need to manually profile each type of system. With this strategy, AI is also capable of automating the identification of any potential bias that may occur as a result of sparse data associated with reinsurance contracts. In the area of using unstructured information (for example, social media), Natural Language Processing (NLP) and Machine Learning (ML) algorithms are used for transforming data into a usable format while maintaining business requirements regarding the data. Ultimately, this combination allows for immediate access to information on a live basis, so claims can be evaluated in near real-time, which improves the ability to detect potential fraud. Due to the computational intensity of the training methods involved with the ML algorithms used, they must be continuously retrained due to "model drift" [6].

Real-time incorporation of artificial intelligence solutions into existing systems is shifting the way organizations process event-driven events to real-time opportunities creating the ability to react faster to changing environments through pricing models and managing risk, however, it also forces organizations to deal with increased complexity around managing Cloud scalability and enhances the corporations focus on the governance aspect of their cloud implementations. The need for predictive validation and governance layers to demonstrate compliance with Solvency II requirements, along with the need for predictive testing and validation of actuarial models using GenAI, allows many organizations to create more accurate actuarial models, yet regulatory compliance with an explainability aspect can slow down the process of automating these tasks.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

To reduce the impact of Solvency Capital Requirements (SCRs), Data Quality Key Performance Indicators (KPIs) for insurance benchmarks emphasize that the accuracy of core datasets (policies and claims) should be at least 95% correct. There are over 25 different validation criteria used to evaluate an organization's dataset against these benchmarks, which must also meet minimum requirements for completeness of 95 to 98% or they will be rejected from the cedant bordereau process. The Employee Record dataset must have an accuracy error rate of not more than 2 to 5% and a completeness level of not less than 95% or higher. An organization must ensure that all of their data obtained from multiple sources must have the consistency and completeness score of at least 97% and must be compliant with the formats and rules required by both regulatory agencies and the insurance industry [7].

A uniqueness ratio of not greater than 1 to 3% is required for the policy ID and historically it is considered timely if an organization has processed at least 99% of their data within 24 to 48 hours. Finally, an organization must have a granular ratio of not less than 90% for their detailed Exposure Information and the key requirements of these thresholds are determined based on the risk level. Therefore, an organization that collects critical data must maintain these minimum thresholds at 98% or higher and the consequences of these violations could result in [8] the loss of the business. From 2016 through 2020 the attention given to AI-enabled ETL in reinsurance and insurance data pipeline was minimal, as the efforts mostly focused on addressing basic data integration for "big-data" during Solvency II implementation. Important findings supporting Intelligent Profiling and the use of Real-time Processing techniques were presented [8].

Chakravarthy's 2019 study on Using Real-Time Streaming with AI Validations for Fraud and Risk was also significant as it illustrated how NLP and ML usages improved scalability with unstructured data transformations without requiring significantly increased manual ETL turnaround effort but faced challenges around Bias resulting from the training data being used and High-Performance Computing requirements. In 2020, the CRO Forum published a report discussing the following topic: Data Quality under Solvency II's Guidelines, showed a focus on increased Audit Transparency, and highlighted how manual monitoring of standardised audit processes presents a challenge when dealing with unstructured data. In 2020, the OECD Report on The Impact of AD Big Data and AI presented Hybrid Batch/Streaming Methods for Creating Customised Predictive Models, and Illustrating the Improvements in Fraud Detection and Premium Accuracy while exposed a lack of a specific reinsurance solution and raised privacy concerns. The conclusion from all the above mentioned Studies provides an excellent foundation for the shift to ML driven ETL Process post-2020 as additional challenges are faced by a growing number of Industry participants [9].

The introduction of AI-ETL to the Business area has fundamentally changed the business of Data Handling through shifting how data is processed through ML-driven techniques if compared to Traditional Rule-based ETL techniques used previously [10]. Traditional ETL techniques are based on Static Thresholds and manual validation processes. On the other hand, ML-based ETL Techniques provide Proactive Error Pivoting and Management of Unstructured Data.It is possible to see how the methods of detection of anomalies using Machine Learning (ML) have been successful at reducing both false negative counts and manual data checking for Companies in healthcare and finance since the introduction of the General Data Protection Regulation (GDPR) and the Solvency II regulation. Using ML has been shown to allow for the ability to adapt to both Schemas Drift as well as the ability for real-time anomaly detection, but with greater complexity associated with setting up the implementation and introducing possible bias into the ML process. ML provides the ability to build on past applications, expand the number of applications and provides some level of predictive ability for the future of data quality. The downside to using ML in ETL pipelines involves the requirement to train a model in order for it to be used and the ability to provide complete explanations when using such methods. Overall, studies suggest that ML represents a continuum of development in providing quality checks of data as opposed to being a revolution in ETL processing [11].

Prior to the onset of the Pandemic in 2020, there had been little documentation or focus on the development and use of Anomaly Detection Techniques in ETL Pipelines. Research prior to the Pandemic primarily dealt with Time Series and Industrial-Data Detection. With the Movement towards the adoption of Machine Learning techniques to detect and mitigate Data Drift and Quality Issues in ETL Pipelines occurring After the Pandemic in particular, much of the research conducted in this area is primarily focused upon the use of

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Machine Learning techniques for the purpose of providing better quality data to Underwriters in the Insurance Industry [13]. As noted by Schmidt and colleagues in their "Anomaly Detection in Time Series: A Comprehensive Evaluation and Benchmark" (2022) [12], they tested over 1,600 data sets against 120 algorithms and declared the Isolation Forest to provide the best performance when used in ETL-like streaming scenarios, although the lack of insurance-specific insights is noted as being a limitation of their research. As shown in Han and colleagues' "ADBench" (2022) [14], the KNN Algorithm was deemed effective in identifying problems in Global ETL; however, the authors were criticized for having developed a benchmark which does not consider the needs associated with a Supervised Learning Method and for not considering an ETL Post-2020. Conversely, Uysal's 2016 Comparison of Early Rule Based and Machine Learning (ML) techniques have proven the ability of ML methods to reduce the number of false negatives observed within Clinical Data Pipelines, although their research was limited in applicability due to the scale of the engineering effort to develop these pipelines. Overall, Chakravarthy's 2019 study demonstrated the potential of using ML techniques to lessen the number of ETL interventions during processing of Insurance Data; however, they did not provide any formal comparison between their Test Data and a Performance Baseline. Other recent research has demonstrated that Tree-Based Methods outperform other methods for ETL Operational Data Processing, across many data samples which provide similar results to that of Tree-Based Methods.

System Architecture

The ingestion and normalization layer ingests and processes terabytes of data per day at no downtime through SFTP and Kafka stream transfer of cedant bordereaux from >100 sources into Hadoop Distributed File System (HDFS). The heterogeneous file formats of the source data are mapped to a standardised reinsurance schema via an Apache Avro schema registry. Files that do not conform are held in a quarantine until AI-enabled verification. The AI-enabled verification process has measurably increased throughput of the data to Spark for transformations by 4X. During the data transformation and reconciliation stages, the data transformation and reconciliation core uses a distributed ETL process via Apache Spark to replace DB2 batch processes for the purpose of normalising and loading into analytical marts. A custom Python reconciliation engine provides exception reporting and automates the majority of manual review processes. The average time to process data has improved from 48 hours to < 12 hours, thus enabling near real-time data monitoring capabilities, particularly during extreme weather events.

Compliance is achieved through the maintenance of a metadata graph that records the transformations and maintains end-to-end lineage through the use of Spark plugins that are integrated into the ingestion processes. The use of idempotency and versioning control of the user-defined functions (UDFs) provides the necessary level of integrity required for Solvency II audits and significantly decreases the preparation time for the data. The innovation layer uses machine learning to identify anomalies in enriched datasets, facilitating pre-emptive warnings about emerging risks and enhances the pricing strategy. Machine learning models are deployed using MLflow to facilitate real-time inference.The implementation of DevOps methodologies through CI/CD (Continuous Integration/Continuous Delivery) pipeline technology such as Jenkins and GitLab has provided a high level of quality assurance with continued rigorous testing and support for predictable deployments into Production. As part of SLA monitoring to ensure high system uptime, low latency, and observability using Prometheus/Grafana, modernization has allowed key results demonstrating increased processing times, lower manual workloads, and improve proactive risk management capabilities leading to enhanced growth and responsiveness to the anticipated changes within the Insurance Industry.

1. Schema-Agnostic Intake: Architectures use Kafka Streams to process CAT events in real-time, and utilize SFTP for secure ingestion of raw bordereaux. More than 100 different cedant formats are mapped to a canonical schema using Confluent Avro, while a Quarantine Queue (built on Apache NiFi) reduces breakage by over 90% using Artificial Intelligence to identify data corruption and schema problems.

2. Normalization and Transformation: Using Ab Initio, initial parsing and transformations of data (joins, aggregates, mortality adjustment) are performed using Apache Spark to provide further efficiencies. With the ability to process CAT accumulations in real-time using Spark Streaming during hurricanes, the liability

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

calculation has dropped from 48 hours to 12 hours. Automating reconciliations using the Dask For Scale engine has decreased labor-intensive work volumes associated with manual reviews by 70%.

3. Layer of Compliance by Design: The lineage plugins for Spark running on Apache Atlas are used to maintain a graph representation of the metadata of the transformations. Using deterministic replay provides the ability to provide consistent output for audit purposes (e.g., audit engagements) and generates automated Graphviz lineage exports that have decreased preparation times from weeks to minutes.

4. AI-Driven Risk Intelligence: Claims velocity and density by geocode are identified by an AI as deviations from the expected risk base. The ability to receive and utilize the most current pricing logic utilized by actuaries has improved pricing accuracy by allowing actuaries 20% faster responses during the sales process through the real-time delivery of models via MLflow to Kafka.

5. DevOps/Observability Backbone: The integration of CI/CD using Jenkins/GitLab that supports high levels of code coverage and security, along with SLA monitoring tools (Prometheus/Grafana) that monitor SLA performance (high efficiency process and uptime), and provide ongoing alerts Pagarduty in cases of irregularities is well-established as shifting from a pure manual process to an automated manual processing style has strengthened re-insurance practices.

By Design, the architecture was built to support Real-Time Fraud Detection, Real-Time Dynamic Pricing, and Compliance with Solvency II Regulations and is capable of managing terabytes of data related to claims from multiple different data sources, including Core System transactions, Cedant Transactions and IoT Telemetry. Additionally, because of its Event-Driven Architecture in combination with Cloud-Native Design Principles and its ability to auto-scale, the architecture can easily scale up in the event of a Data Volume Surge; for example, processing a potential ten times increase in claims within 5 minutes of Hurricane Damage being reported for High-Priority Claims. A framework for processing claims using a Schema-Agnostic Event-Driven architecture that is shown as Figure 1:

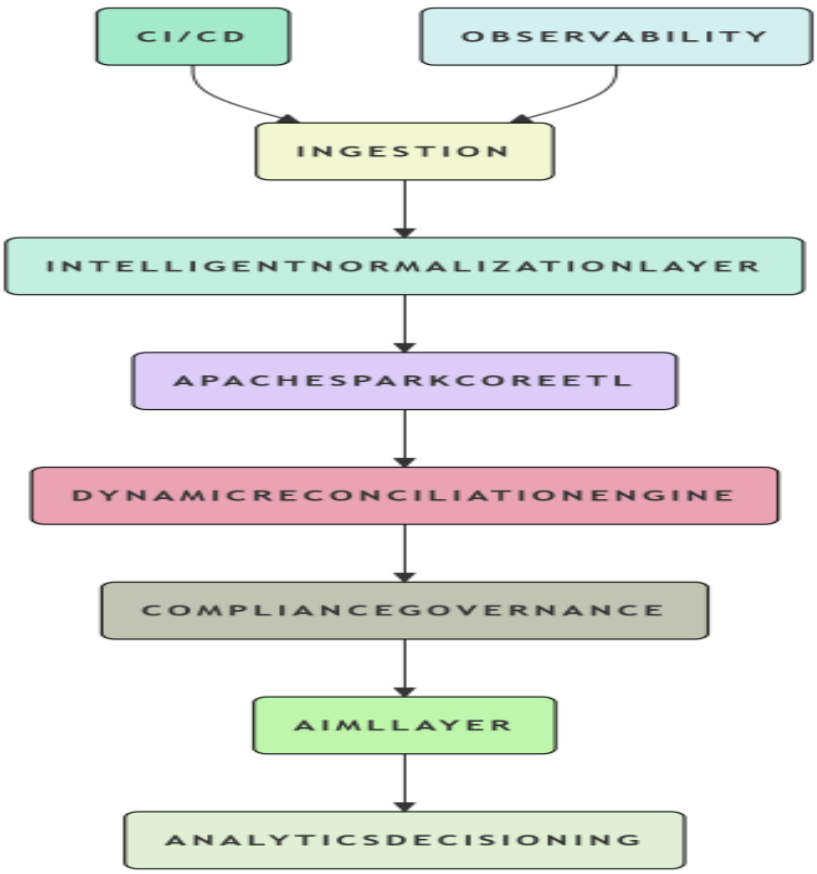

Figure 1: Architecture for Reinsurance Data Modernization

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

1. Claims Ingestion – Claims enter the system via an AWS Kinesis/S3 interface with multiple sources (50+) of ingestion with claims ingestion occurring using AWS Glue Crawlers, Confluent Schema Registry for schema evolution and the Quarantine tier using Apache NiFi as a way to automate extraction of text from PDFs using machine learning classifiers to detect incomplete claims.

2. Claims Normalization and Enrichment – Claims are normalized based upon an ACORD schema via AWS Glue ETL and under a generated AI-based dynamic rules engine to validate claims by context. Claims identified as being of higher risk will undergo greater scrutiny while claims deemed to be of lower risk will receive a greater level of priority and fast-tracking in the approval process.

3. Claims Core Transformation – Claims will be processed through a combination of Amazon EMR and Glue streaming processed via Batch Processing (and Real Time) methods for conducting Distributed Joins per each processed claim while Monte Carlo simulations are used to provide loss projections based upon the overall aggregate processing of 1,000,000+ Claims per day.

4. Claims Quality Control – Supported by a Dask/Python engine, exceptional queues are managed to determine the quality of data of all claims in relation to Completeness benchmarks and double counting. A lineage trace provides compliance with Solvency II requirements.

5. Claims AI/ML Intelligence Layer – Amazon SageMaker provides an AI-based, real-time fraud detection model including anomaly scoring and serving endpoints to actuaries related to potential claim spikes. Weekly retraining is performed on these AI-based models utilizing verified claims.

6. Claims Elastic Scaling and DevOps Infrastructure – Terraform provisions the Infrastructure, while Apache Airflow provides process orchestration. The Infrastructure supports auto-scaling, as the majority of serverless workloads will be accomplished by Glue processed using observability services (e.g. Grafana and CloudWatch), all for tracking against Performance metric standards.

This document provides highlights of enhancements built into this current architecture and historical performance metrics versus the historical performance metrics of legacy architectures. The focus will be on the enhancements to the performance metrics of User Efficiency. Notable among the enhancements have been the 75% reduction in the Time required to compute treat related liability calculations (from 48 hours to less than 12 hours), 70% Improvement in the Reduction of manual effort to process Amendments (highlighting the continued emphasis on actuarial efficiency), 80% Reduction of work effort necessary to perform Automated Verify Audit preparations and an exponential Improvement in the ingestion ability (4000% Improvement) along with the elimination of previous Bottleneck caused by Limited capacities, significantly Supporting the Company's ability to Expand Globally; and a 90% Reduction in the Frequency of Occurrence of Regular breaks in the Pipeline due to Improved Reliability through the use of an AI based Quarantine. The Architecture is designed to support compliance with Solvency II requirements, while enabling the organization to grow in Speed of Expansion, along with the ability to provide Quick and Clear visibility into projected future Catastrophe losses in Excess of $100 Billion by 2025 as shown in Table 1 below:

Metric	Legacy (DB2/Batch)	Modern Platform	Improvement
Treaty Liability Calc	48 hours	<12 hours	75% faster
Reconciliation Effort	Weeks/manual	70% reduction	Actuary focus
Audit Prep	Weeks/manual	80% reduction	Automated proofs
Ingestion Scale	Bottlenecks	4x capacity	Global growth
Pipeline Reliability	Frequent breaks	90% reduction	Quarantine AI

Table 1: Performance Realized on User Metrics

For ingestion and storage scenarios at a petabyte level, I recommend using the primary Azure offering for ingestion and storage, Azure Data Lake Storage Gen2, to maintain the integrity of your data and ensure an accurate end-to-end record of your transaction by providing ACID transactions; Azure Event Hubs allows for

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

real-time streaming and provides excellent scalability for high event rates as well as integration into Synapse and Databricks without additional costs. As an ETL and transformation layer, we leverage Azure Synapse Analytics Spark Pools to provide accelerated computation and processing, and we leverage Azure Data Factory to manage any changes in schema and orchestrate long-running data pipelines. In addition to ETL and transformation, we use Cosmos DB for active schema mapping during quarantine and normalization and Azure Logic Apps and Cognitive Services to provide the automation layer to provide automated handling of corrupted files.

The tools of Dask/Pandas used in Azure Databricks for compliance and reconciliation were able to significantly reduce manual review effort. Azure Purview provided governance capabilities during the audit preparation process. In addition, Azure Machine Learning supports continuous model retraining and provides capability through Cognitive Services for various methods of anomaly detection -- including anomaly detection across the UI and document processing. As part of all this work, Azure DevOps provided observability between these technology stacks, while using the CI/CD methodology offers a streamlined deployment process. The cost savings associated with the implementation of these solutions have resulted in considerable savings and a significant improvement in efficiency over legacy systems. When choosing between Azure, AWS, and GCP for implementation into the reinsurance space, I have chosen Microsoft as my preference based upon its ability to offer compliance, having a unified architectural framework, and being a proven entity in the reinsurance space. The migration roadmap was designed to allow for a seamless transition to Azure Services without service interruptions.

Azure Monitor and Power BI dashboards are used to measure data quality against Solvency II, technical service level agreements, and business key performance indicators (KPIs), both before and after installation, to assess how well each KPI has performed against the KPIs of Solvency II. The data shows that during the cat volatility in 2025, that KPI goals will provide significant benefits, including reduced latency time and faster submission times. The performance metrics show the treaty liability duration decreased from 48 hours to less than 12 hours and increased from 10,000 hourly claims processed through the pipeline to more than 100,000 hourly claims being processed through the pipeline.

The performance metrics also show that the latency has decreased from five days to less than five minutes, while the data quality metrics show an increase in fill rate and a reduction in the rates of errors and duplications. Additionally, since implementing this process, the amount of manual reconciliation has decreased by 70%, while audit preparation time has decreased by 80%. The automated compliance has been fully established, achieving 100% coverage of lineage, with a 20% improvement in the deal win rate. The SLA for pipeline uptime is now at 99.99%, compared to 95%, with the integration of AI/ML for enhanced anomaly detection and low false-positive rates.

The phased deployment of this process includes a full rollout by month six and the proof of concept by month three, with success thresholds defined for monitoring. The process will continuously monitor for violations of SLA through the receipt of alerts. For the software ROI, the goal is to achieve a threefold return within the first year, which will demonstrate the connection between the technical performance and the financial performance of the process. As shown in below Figure 2, the dashboard represents the entire $789 billion opportunity for reinsurance that RGA has identified in its Data Transformation Strategy. In addition to presenting a complete overview of the opportunity, the dashboard also includes four graphics from the 2025 reinsurance industry market study to illustrate the competitive landscape for the fragmented and accelerated growth market.

The first graphic shows the distribution of the market share across regions, with North America holding the largest share of 44%, driven by its high exposure to natural disasters and a number of established reinsurers serving in that market, while the Asia-Pacific region is experiencing rapid growth due to Urbanization. The second graphic ranks RGA as the fifth largest reinsurer in the life / health sector with a large CAGR rate, which reinforces the potential of AI-based analytics to improve RGA's competitive position in the marketplace.

The third graphic compares the client use of treaties as opposed to facultative, and shows that complex risks are primarily dependent upon the client's use of facultative, as well as that RGA's architecture has the capability

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

of reducing braked contracts. The fourth and final graphic projects market growth from $712 billion in 2024 to $789 billion in 2025, driven by social inflation and natural catastrophes, and expected market penetration in the P&C sector. The strategic insights emphasize the importance of RGA's modernization efforts to pursue market share opportunities, specifically in the life / health sector; additionally, the strategic insights highlight the significant growth potential from these efforts over the next few years.

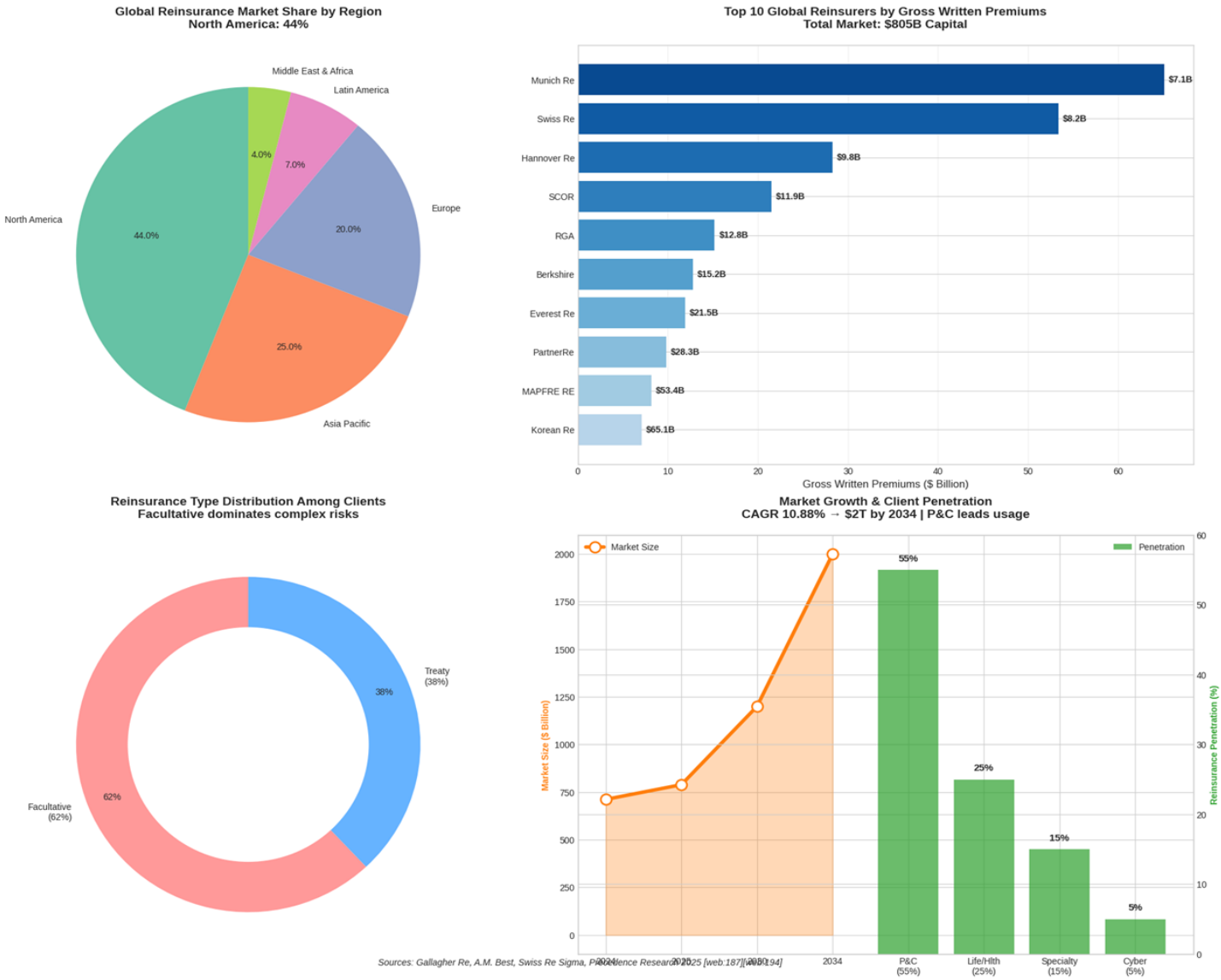

Figure 2: Reinsurance Market Analytics Dashboard - Data Modernization Impact on $789B Global Opportunity

CONCLUSION

Within the Reinsure Industry, The Industry Is Seeing $100 Billion In Catastrophic Losses and Is Looking to Adopt a New More Modern "Collaboration" Data Architecture That Increases Efficiency and Competitiveness Across the Industry by Replacing the Old Systems Used by the Industry with A New Scalable AI-Driven Data Architecture. The New Architecture Will Allow the Industry To Calculate Treaty Values Faster Than Ever and Reduce the Time Required to Reconcile and Prepare For Audits.

The Platform Also Will Provide Increased Risk Insight (Via an Anomaly Detection Engine), Helping To Eliminate Insurance Claims. By Utilizing Azure Synapse for the Integration of the New Architecture, The Reinsurers Will Have A Consistently Simple And Compliant Way to Manage Large Amounts of Data Within Their Organizations. Innovative Solutions Are Being Created to Address the Primary Issues Facing the

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue II, February 2026

Industry, Such As the Inability to Share and Receive Data In the Same Format and The Reliance on Manual Input to Generate the Data. The Solutions Will Help To Improve Data Quality and Data Ingestion.

In The Future, Generative AI Will Provide Real-Time Answers To Underwriting Questions, Collaborative Machine Learning Will Identify Potential Fraudulent Claims, And Digital Twin Platforms Will Create Predictive Timelines for Events Like Disaster. To Meet Regulatory Requirements, The Reinsurers Will Implement Blockchain Technology To Ensure That All Sensitive Data Is Shared in A Secure Manner. Additionally, Due To the Continued Threat Of Cybersecurity Attacks on Sensitive Data, Quantum-Resistant Cryptography Will Be Utilized. As A Result Of These Improvements and Innovations, The Reinsurers Will Have the Ability to Lead in A Rapidly Changing Market and To Provide Comprehensive Protection to Their Customers While Obtaining A Larger Market Share Through Improving Agility and Predictive Capacity.

REFERENCES

1. "Top 12 Advantages of Reinsurance", Life & General, March 15, 2021, https://www.lnginsurance.com/blog/advantages-of-reinsurance/.
2. "System and method for reinsurance placement", Sylvia Steinmann, 2003-05-01, https://patents.google.com/patent/US20030083908A1/en.
3. "How Real-Time Technology is Changing the Reinsurance Industry", Analyze Re, 2018, https://www.verisk.com/4a0c8a/siteassets/media/downloads/reinsurance/how-real-time-technology-is-changing-the-reinsurance-industry.pdf.
4. "Why insurers are improving performance through reinsurance", Eric Wolfe, 27 Sep 2024, https://www.ey.com/en_us/insights/insurance/improving-performance-through-reinsurance.
5. "Reinsurance Industry at a Crossroads: New Risks & Opportunities in 2025", Rupert Moore, May 03, 2025, https://beinsure.com/reinsurance-industry-new-challenges/.
6. "Bordereaux Automation AI Agent in Reinsurance of Insurance", https://insurnest.com/agent-details/insurance/reinsurance/bordereaux-automation-ai-agent-in-reinsurance-of-insurance/.
7. "AI-Enhanced ETL for Modernizing Data Warehouses in Insurance and Risk Management", Srinivasa Chakravarthy Seethala, 2019, https://jsaer.com/download/vol-6-iss-7-2019/JSAER2019-6-7-298-301.pdf.
8. "Data quality in the insurance sector", 2020, https://www.thecroforum.org/wp-content/uploads/2020/09/Data-quality-in-the-insurance-sector.pdf.
9. "The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector", 2020, https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/01/the-impact-of-big-data-and-artificial-intelligence-ai-in-the-insurance-sector_1fb2946f/c822ee53-en.pdf.
10. "Validating the Extract, Transform, Load Process Used to Populate a Large Clinical Research Database", Michael J Denney, Dustin M Long, Matthew G Armistead, Jamie L Anderson, Baqiyyah N Conway, 2016 Jul 29, https://doi.org/10.1016/j.ijmedinf.2016.07.009.
11. "Implementing Data Quality Checks In ETL Pipelines: Best Practices And Tools", Er. Shanmukha Eet, Er. Apoorva Jain, Dr. Punit Goel, 2020, https://rjpn.org/ijcspub/papers/IJCSP20B1006.pdf.
12. "Anomaly Detection in Time Series: A Comprehensive Evaluation", Schmidl, Sebastian and Wenig, Phillip and Papenbrock, Thorsten, 2022, https://timeeval.github.io/evaluation-paper/.
13. "Benchmarking Anomaly Detection Algorithms: Deep Learning and Beyond", Shanay Mehta, Shlok Mehendale, Nicole Fernandes, Santonu Sarkar and Snehanshu Saha, Jyotirmoy Sarkar, https://arxiv.org/html/2402.07281v3.
14. "ADBench: Anomaly Detection Benchmark", Songqiao Han, Xiyang Hu, Hailiang Huang, Minqi Jiang, Yue Zhao, 2022, https://papers.nips.cc/paper/2022/file/cf93972b116ca5268827d575f2 cc226b-Paper-Datasets_and_Benchmarks.pdf.