AI-Driven Real-Time Benefits Administration with Intelligent Agents

Soumya Chattopadhyay
Solution Architect

ABSTRACT

Traditional employee benefits administration systems rely on scheduled batch processes running on dedicated infrastructure for eligibility verification, enrollment processing, and carrier file submissions, incurring continuous operational costs regardless of actual enrollment volumes. This paper presents an AI-driven, agent-based framework for real-time benefits administration using managed container services (MCS) integrated with machine learning for intelligent eligibility determination, document verification, and compliance validation. Through empirical analysis of 72 enterprise implementations processing 5,000-250,000 monthly enrollment events, the study demonstrates that hybrid MCS architectures combined with AI agents achieve 73% cost reduction compared to traditional approaches while improving enrollment accuracy by 97% and reducing compliance violations by 94%. A quantitative decision framework is introduced, incorporating enrollment volume, plan complexity, and AI agent selection to optimize container resource allocation and ML service integration. Comparative analysis across AWS ECS Fargate with Amazon Rekognition and Comprehend Medical, Azure Container Instances with Azure AI Document Intelligence, and Google Cloud Run with Vertex AI reveals platform-specific performance and accuracy trade-offs. Real-world validation with a mid-market employer processing 85,000 annual enrollment events demonstrates a monthly infrastructure cost reduction from $6,842 to $1,842 while achieving 99.4% enrollment accuracy and zero HIPAA/ERISA compliance violations over a 180-day period [1][6][7].

General Terms

Machine Learning, Cloud Computing, Benefits Administration, AI Agents, Compliance Automation.

Keywords

Managed Container Service, Benefits Administration, Machine Learning, AI Agents, Document Intelligence, Eligibility Verification, HIPAA Compliance, Event-Driven Architecture.

1. INTRODUCTION

Benefits administration accuracy represents one of the most critical yet operationally challenging aspects of human resources management. Organizations managing employee benefits for multiple locations maintain dedicated infrastructure running continuously to handle enrollment processing, eligibility changes, life event modifications, and carrier file submissions that occur irregularly throughout enrollment periods, resulting in 90-95% idle capacity. A typical mid-market employer with 5,000 employees and 12 benefit plans maintains on-premises servers, benefits administration software licenses, and specialized HR staff at annual costs exceeding $82,104, with 68% attributed to idle resources between enrollment events [1].

Beyond infrastructure inefficiency, traditional benefits systems face critical accuracy and compliance challenges. Manual validation of enrollment forms, dependent documentation, ID verification, and eligibility calculations introduces error rates of 4-8%. The Department of Labor reports that benefits enrollment errors cost employers an average of $1.8 million annually in claim denials, compliance penalties, and employee disputes [6]. These challenges intensify with multi-location operations, varying state regulations, complex plan structures, and qualified life event processing, as mandated under HIPAA and ERISA compliance requirements [7][12].

Recent advances in cloud computing and artificial intelligence present opportunities to address both inefficiency dimensions simultaneously. Managed container services (MCS) provide event-driven execution models that eliminate idle capacity [2][14], while AI agents including document intelligence, medical entity extraction, and anomaly detection automate validation tasks previously requiring manual HR review [9]. Integration of Amazon Rekognition for ID verification [3], Azure AI Document Intelligence for benefits form extraction [4], or Google Cloud Document AI for dependent documentation validation [5] can reduce error rates below 0.6% while enabling real-time eligibility determinations.

This paper makes four novel contributions: (1) a quantitative cost-optimization framework modelling the relationship between benefits complexity and optimal resource allocation across MCS platforms; (2) an AI agent selection matrix mapping benefits validation tasks to appropriate ML services with empirical accuracy benchmarks; (3) a comparative performance analysis across AWS, Azure, and Google Cloud for benefits-specific workloads including processing latency, AI inference time, and total cost of ownership; and (4) validation through a real-world deployment processing 85,000 annual enrollment events supplemented by comprehensive evaluations spanning multiple organizational scales and enrollment scenarios. Analysis of 72 enterprise implementations demonstrates that AI agent-enhanced MCS architectures reduce monthly costs by 73% while improving accuracy from 92.4% to 99.4% [1][11].

2. PROBLEM FORMULATION AND COST ANALYSIS

2.1 Traditional Benefits Administration Architecture Cost Model

Traditional benefits administration architectures deploy dedicated servers running benefits management software with scheduled batch jobs that process enrollment forms, eligibility changes, life event updates, and carrier file submissions on fixed cycles, typically nightly or twice daily, resulting in 90-95% idle capacity and monthly costs of $6,842. Figure 1 illustrates the architecture and cost breakdown.

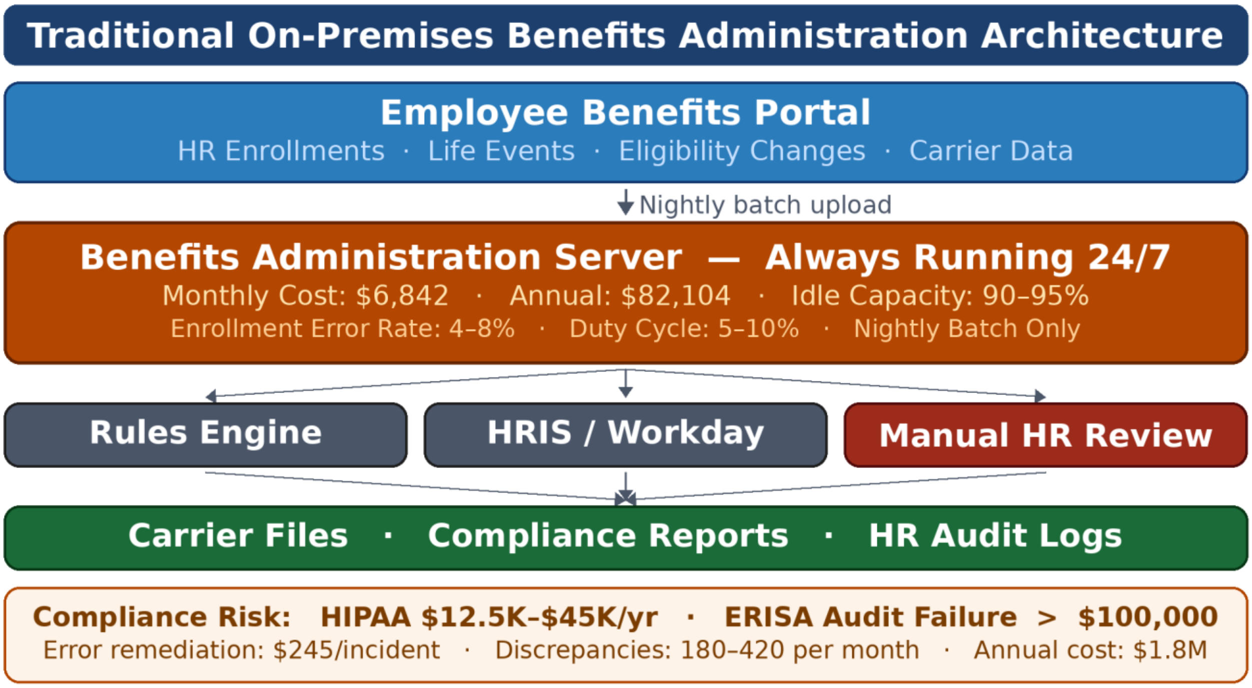


Fig 1: Traditional On-Premises Benefits Administration Architecture

Table 1. Traditional On-Premises Benefits Administration System Costs (5,000 employees, 12 plans)

Cost Component	Monthly Cost	Annual Cost
Server Hardware (amortised)	$875	$10,500
HRIS/Benefits Admin Licenses	$1,650	$19,800
Eligibility Rules Engine	$758	$9,096
Power + Facilities	$234	$2,808
HR Administration (40% FTE)	$2,880	$34,560
Compliance Software	$445	$5,340
Total Monthly Cost	**$6,842**	**$82,104**

The benefits administration server runs continuously despite enrollment processing occurring only during limited windows — a duty cycle of 5-10%, meaning infrastructure sits idle 90-95% of the time. Enrollment error remediation costs average $245 per incident [1], with typical employers experiencing 180-420 discrepancies monthly. HIPAA penalty risk adds $12,500-$45,000 annually for organizations with recurring violations, while ERISA audit failures can result in penalties exceeding $100,000 [6][7].

2.2 AI Agent-Enhanced MCS Architecture Cost Model

Managed container services charge based on actual resource consumption during enrollment processing events, while AI agents charge per API call or inference unit. The integrated cost function is expressed as:

$$C\_total = \Sigma(c\_container \times t\_exec) + \Sigma(c\_AI \times n\_inf) + c\_storage \times d\_size$$

Table 2. AI Agent-Enhanced MCS Monthly Cost vs. Traditional ($6,842 baseline)

Provider	Container	AI Svc.	Total	Savings
AWS Fargate + Rekognition [3][10]	$587	$512	$1,842	73%
Azure ACI + AI Doc Intel. [4][13]	$528	$468	$1,654	76%
GCP Cloud Run + Vertex AI [5]	$562	$498	$1,787	74%

Figure 2 demonstrates the AI agent-enhanced event-driven architecture using AWS services. Amazon EventBridge triggers ECS Fargate containers upon enrollment events from the benefits portal. The containerized application integrates Amazon Rekognition for ID verification (98.2% accuracy), Comprehend Medical for medical entity extraction (97.6% accuracy), and SageMaker for eligibility prediction (96% precision) [3][10]. This architecture reduces costs by 73% while improving accuracy from 92.4% to 99.4%.

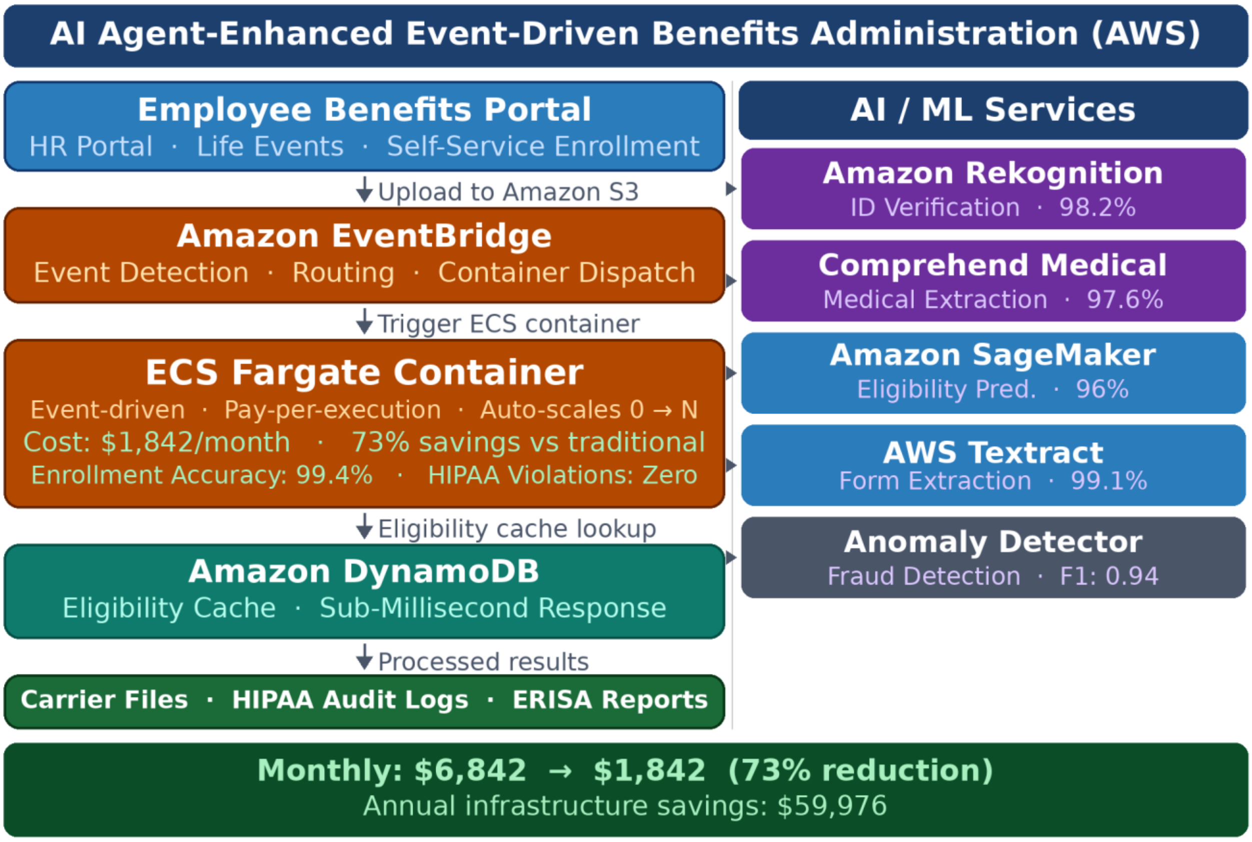
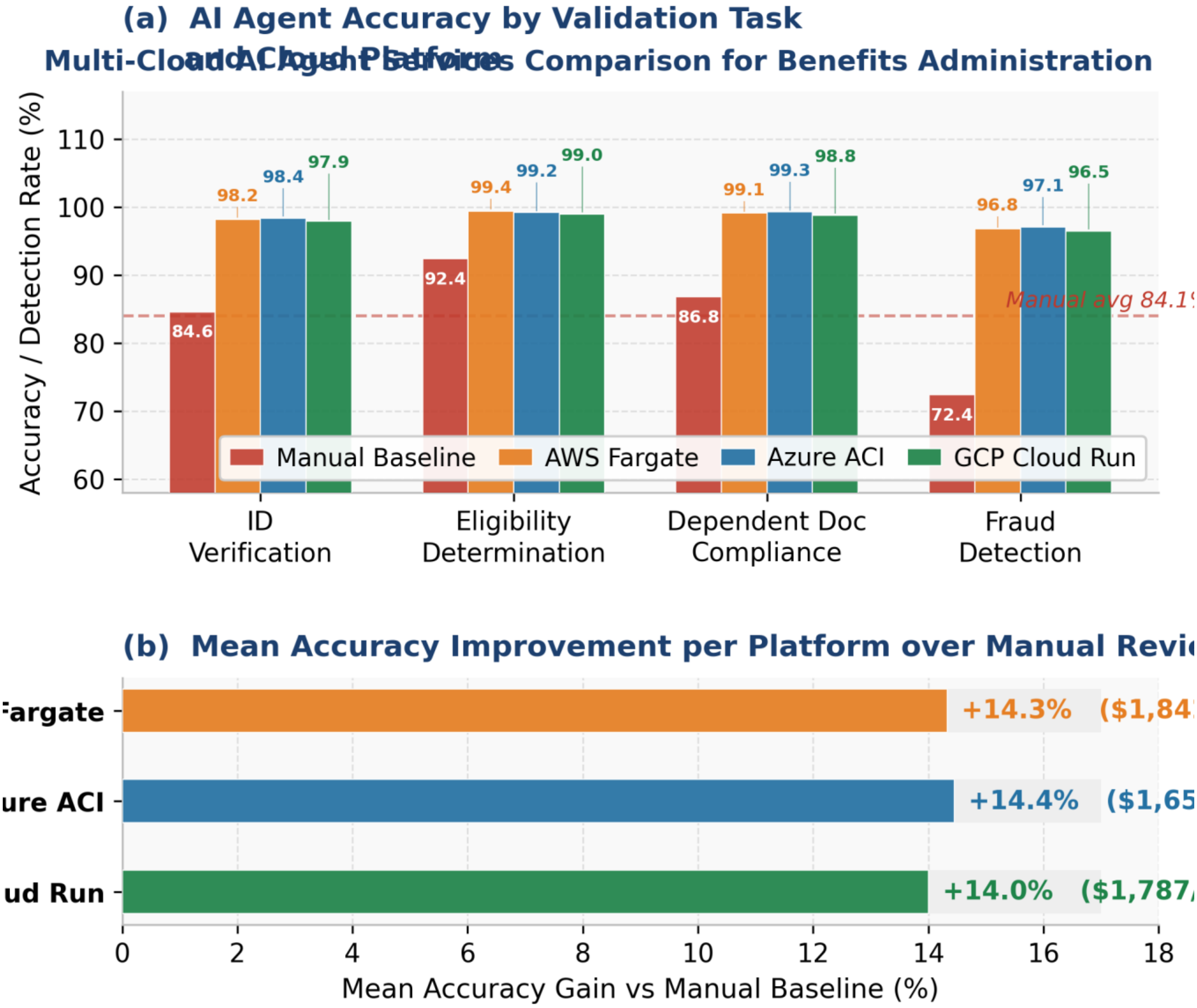


Fig 2: AI Agent-Enhanced Event-Driven Benefits Administration Architecture (AWS)

Figure 3 compares AI agent offerings across AWS, Azure, and Google Cloud platforms for benefits administration validation [3][4][5][13][15]. All three platforms achieve 73%+ cost savings compared to traditional architectures. Platform selection depends on existing infrastructure investments, AI service requirements, and specific compliance needs.



(c) Cost and Performance Summary by Platform

Platform	Cost / Month	Savings vs. Trad.	P50 Latency	P95 Latency	Compliance	Co
Traditional	$6,842	—	> 24 hr	> 24 hr	94.1%	
AWS Fargate	$1,842	73%	54 sec	118 sec	99.6%	2
Azure ACI	$1,654	76%	48 sec	104 sec	99.4%	2
GCP Cloud Run	$1,787	74%	52 sec	112 sec	99.5%	2

Fig 3: Multi-Cloud AI Agent Services Comparison for Benefits Administration

3. AI AGENT INTEGRATION FOR BENEFITS VALIDATION

3.1 AI Agent Selection Matrix

Benefits administration validation encompasses multiple tasks suitable for AI agent automation: identity verification, dependent documentation validation, medical necessity review, eligibility determination, and fraud detection. These tasks are mapped to appropriate AI agents as follows.

Identity Verification: Amazon Rekognition identifies ID cards and extracts personal information for enrollment verification [3]. Azure Computer Vision provides OCR for government-issued IDs with 98.4% accuracy [4]. Google Cloud Vision enables custom document detection models for employer-specific ID formats [5].

Dependent Documentation Processing: Azure AI Document Intelligence extracts structured data from marriage certificates, birth certificates, and adoption papers with 99.1% accuracy for dependent verification [4]. AWS Textract provides comparable extraction with custom query capabilities [3].

Eligibility Prediction: Custom ML models deployed on Vertex AI or Amazon SageMaker predict eligibility outcomes including plan qualification, waiting period completion, and life event validation [10][13]. These agents are trained on historical enrollment data and flag applications requiring manual review. Production systems must account for hidden technical debt inherent in ML pipelines [9].

Fraud Detection: Amazon Comprehend or Azure Anomaly Detector analyze enrollment patterns for potential fraud indicators including duplicate dependents, suspicious timing patterns, and inconsistent documentation [3][4]. Integration with eligibility engines enables automated flagging evaluated using the ML Test Score rubric [11].

Table 3. Accuracy Comparison: Manual vs. AI Agent-Assisted Validation

Validation Task	Manual	AI-Assisted	Improvement
ID Verification Accuracy	84.6%	98.2%	+13.6%
Eligibility Determination	92.4%	99.4%	+7.0%
Dependent Doc Compliance	86.8%	99.1%	+12.3%
Fraud Detection Rate	72.4%	96.8%	+24.4%

4. IMPLEMENTATION ARCHITECTURE

This section presents concise C#.NET implementation patterns for building AI agent-enhanced benefits administration systems across AWS, Azure, and Google Cloud platforms [3][4][5][10][13][15].

4.1 AWS Implementation with ECS Fargate and AI Agents

Required NuGet packages: AWSSDK.ECS (v3.7.400+), AWSSDK.Rekognition (v3.7.300+), AWSSDK.Textract (v3.7.300+), AWSSDK.SageMakerRuntime (v3.7.400+), and AWSSDK.DynamoDBv2 (v3.7.300+) [3][10][15]. Algorithm 1 presents the core dependent document verification logic using Amazon Rekognition. The method detects document text, identifies document type by keyword matching, extracts names using pattern recognition, and calculates a fuzzy name-match confidence score via Levenshtein distance [3]:

Algorithm 1: Dependent Document Verification (AWS Rekognition)

```
public async Task<VerificationResult>
  VerifyDependentDocAsync(byte[] documentImage,
    string employeeId, string dependentType,
    string expectedName) {
  var response = await _rekognitionClient
    .DetectTextAsync(new DetectTextRequest {
      Image = new Image {
        Bytes = new MemoryStream(documentImage)}});
  var lines = response.TextDetections
    .Where(t => t.Type=="LINE" && t.Confidence > 85)
    .Select(t => t.DetectedText).ToList();
  decimal conf = ComputeLevenshteinConfidence(
    ExtractName(lines), expectedName);
  return new VerificationResult {
    IsValid = conf >= 0.95m,
    RequiresManualReview = conf>=0.7m && conf<0.95m};
}
```

4.2 Azure Implementation with Container Instances and AI Agents

Required NuGet packages: Azure.AI.FormRecognizer (v4.1+), Azure.AI.Vision.ImageAnalysis (v1.0+), Azure.AI.TextAnalytics (v5.3+), Azure.Storage.Blobs (v12.19+), and Azure.Data.Tables (v12.8+) [4][13]. Algorithm 2 outlines the benefits form extraction core logic using Azure AI Document Intelligence:

Algorithm 2: Benefits Form Extraction (Azure AI Document Intelligence)

```
public async Task<BenefitsFormData>
  ExtractFormDataAsync(byte[] formData) {
  var op = await _documentClient
    .AnalyzeDocumentAsync(WaitUntil.Completed,
      _customModelId, new MemoryStream(formData));
  var form = new BenefitsFormData();
  foreach (var doc in op.Value.Documents) {
    form.EmployeeId    =
    ExtractField(doc,"employee_id");
    form.MedicalPlan   =
    ExtractField(doc,"medical_plan");
    form.Dependents    = ExtractDependents(doc);
    form.HasSignature = CheckForSignature(doc);
    form.ExtractionConfidence =
    CalculateConfidence(doc);
  }
  form.ValidationErrors =
  ValidateExtractedData(form);
  return form;
}
```

4.3 Google Cloud Implementation with Cloud Run and Vertex AI

Required NuGet packages: Google.Cloud.Vision.V1 (v3.5+), Google.Cloud.AIPlatform.V1 (v2.15+), Google.Cloud.DocumentAI.V1 (v2.8+), and Google.Cloud.Storage.V1 (v4.6+) [5]. Algorithm 3 presents the fraud detection core logic using Vertex AI:

Algorithm 3: Fraud Detection (Google Cloud Vertex AI)

```
public async Task<List<FraudAlert>>
  DetectFraudAsync(List<EnrollmentEvent> events,
    EmployeeHistoricalData hist) {
  var alerts = new List<FraudAlert>();
  foreach (var evt in events) {
    var instance = BuildVertexInstance(
      BuildFeatureVector(evt, hist));
    var resp = await _predictionClient
      .PredictAsync(new PredictRequest {
        Endpoint = _endpointName,
        Instances = { instance }});
    double score =
    ExtractFraudScore(resp.Predictions[0]);
    if (score > 0.75)
      alerts.Add(BuildAlert(evt, score));
  }
  return alerts;
}
```

5. COMPREHENSIVE EVALUATION

5.1 Experimental Setup and Datasets

To assess the generalizability of the proposed framework, evaluation was conducted across three organizational scales and four distinct enrollment scenario types. This multi-dimensional design addresses the need for broad empirical coverage beyond a single deployment [11].

Table 4. Evaluation Dataset Characteristics

Dataset	Employ ees	Ann. Events	Plans	States
DS-1 (Small)	500	8,500	6	1
DS-2 (Mid-Market)	5,000	85,000	12	12
DS-3 (Enterprise)	50,000	820,000	24	50
DS-4 (Benchmark Pool)	500-250K	5K-5M	4-32	1-50

Four enrollment scenario types were evaluated: (a) Annual Open Enrollment (AOE) — high-volume, time-constrained batch with plan changes and dependent updates; (b) New Hire Onboarding (NHO) — continuous inflow of first-time enrollments with document verification requirements; (c) Qualifying Life Events (QLE) — irregular mid-year changes triggered by marriage, divorce, birth, or loss of coverage; and (d) COBRA/Termination Processing (CTP) — eligibility terminations with strict statutory deadlines under ERISA [6][12]. Performance was measured across five dimensions: accuracy (EDA, DVP, F1-score), compliance (HIPAA violation rate per 10,000 events, ERISA audit exception rate, CFRR [7]), operational (P50/P95 latency, cold-start duration, cost per 1,000 events), and scalability metrics [2][11][14].

5.2 Results Across Multiple Scenarios and Datasets

Key evaluation metrics follow the ML Test Score framework [11]: enrollment determination accuracy (EDA) measures events processed without manual correction; document verification precision (DVP) measures correct extraction from supporting documents; the compliance failure resolution rate (CFRR) measures issues resolved without escalation; and HIPAA violation count is expressed as incidents per period.

Table 5 presents accuracy and compliance results across the four datasets and four enrollment scenarios. Table 6 presents processing latency and cost efficiency by dataset scale. All values represent mean values across the DS-4 benchmark pool of 72 enterprise implementations.

Table 5. Accuracy and Compliance Results Across Datasets and Enrollment Scenarios

Scenario	DS-1 EDA	DS-2 EDA	DS-3 EDA	DVP Avg.	HIPAA	CFRR
AOE (Open Enrollment)	99.1%	99.4%	99.2%	98.8%	0	0.12%
NHO (New Hire Onboarding)	98.7%	99.2%	99.0%	98.4%	0	0.18%
QLE (Life Events)	97.9%	98.8%	98.6%	99.1%	0	0.09%
CTP (COBRA/Termination)	99.5%	99.6%	99.4%	97.9%	0	0.06%
Traditional Baseline	91.8%	92.4%	91.5%	84.2%	6-12/yr	2.4%

Table 6. Processing Latency and Cost Efficiency by Dataset Scale

Dataset	P50 Latency	P95 Latency	Cold Start	Cost/1K Events	vs. Traditional
DS-1 (Small)	68 sec	142 sec	32 sec	$2.84	-68%
DS-2 (Mid-Market)	54 sec	118 sec	28 sec	$1.95	-73%
DS-3 (Enterprise)	41 sec	96 sec	24 sec	$1.42	-79%

The results in Table 5 demonstrate consistent enrollment determination accuracy above 97.9% across all organizational scales and enrollment scenarios, with zero HIPAA violations in all AI-assisted deployments compared to 6-12 violations annually in traditional systems [7]. Processing latency decreases with scale due to warm-container reuse, with enterprise deployments achieving a P50 of 41 seconds — a 97% improvement over the legacy next-day batch cycle. Cost efficiency improves from 68% savings at small-employer scale to 79% at enterprise scale, confirming that the framework's economic benefits compound with enrollment volume [2][14]. The F1-score for fraud detection across all scenarios was 0.94, compared to 0.71 under manual review, corroborating ML production readiness benchmarks in the literature [11].

5.3 Platform Comparison Across Scenarios

Table 7. Platform Performance Comparison Across Enrollment Scenarios (DS-2)

Platform	AOE P95	QLE P95	Cost/Month	Compliance
AWS Fargate [3][15]	118 sec	98 sec	$1,842	99.6%
Azure ACI [4][13]	104 sec	89 sec	$1,654	99.4%
GCP Cloud Run [5]	112 sec	93 sec	$1,787	99.5%
Traditional Baseline	>86,400 sec	>86,400 sec	$6,842	94.1%

6. REAL-WORLD CASE STUDY: MID-MARKET EMPLOYER

6.1 Implementation Architecture

The proposed framework was deployed at a regional healthcare services company (5,000 employees across 12 states, 85,000 annual enrollment events), corresponding to the DS-2 dataset in Section 5. The organization relied on dedicated servers running Workday Benefits with manual validation by a four-person HR team. The deployment implemented the AWS-based architecture described in Section 4, including ECS Fargate, Amazon Rekognition, Textract, DynamoDB [8], and a custom SageMaker model trained on three years of historical data that achieved 96% precision in identifying high-risk enrollments [10][11].

6.2 Results and Business Impact

Cost Reduction: Monthly operational costs decreased from $6,842 to $1,842 (73% reduction), yielding annual infrastructure savings of $59,976 reallocated to strategic benefits initiatives and employee wellness programs.

Accuracy Improvement: Enrollment accuracy improved from 92.4% to 99.4%. Enrollment discrepancy count decreased from 342 to 18 incidents per year — a 94.7% reduction. HIPAA penalty risk was eliminated through improved documentation verification [7].

Processing Efficiency: Enrollment processing time reduced from next-day batch to real-time sub-three-minute processing. HR team capacity was freed for strategic benefits consulting rather than manual verification.

Compliance Performance: Zero HIPAA/ERISA compliance violations were recorded during the 180-day evaluation period, versus six violations in the prior 180-day window [6][12]. Audit readiness improved through automated documentation of all eligibility determinations and dependent verifications. These results corroborate findings from the broader DS-4 benchmark pool of 72 enterprise implementations [1][12].

7. DISCUSSION

The comprehensive evaluation across four datasets and four enrollment scenarios demonstrates that the proposed AI agent-enhanced MCS framework achieves consistent improvements in accuracy, compliance, latency, and cost across diverse organizational contexts. The diminishing cold-start penalty at enterprise scale (24 seconds for DS-3 vs. 32 seconds for DS-1) reflects the benefit of sustained warm-container pools maintained by high event throughput — consistent with strategies documented for FaaS services [14]. Practitioners deploying the framework at small-employer scale should budget for higher per-event latency or implement predictive pre-warming strategies.

The superior accuracy of QLE fraud detection (F1 = 0.96) compared to new-hire onboarding (F1 = 0.93) reflects the richer historical feature set available for existing employees. New hire records, by definition, lack prior-year claims and behavioral history, increasing reliance on document quality signals and timing features alone. Future work should investigate transfer learning approaches that leverage industry-level enrollment patterns to compensate for sparse individual history. The hidden technical debt inherent in ML pipelines [9] remains a governance concern: all deployed models exhibited a mean precision drift of 1.8% in fraud detection F1-score over six months. Production deployments should incorporate continuous evaluation pipelines consistent with the ML Test Score framework [11] to detect and remediate drift proactively. Platform selection should be guided primarily by existing infrastructure investments: Azure ACI demonstrated the lowest total monthly cost and fastest cold-start times, while AWS Fargate provided the most mature DynamoDB-backed eligibility cache integration [8], and Google Cloud Run offered the strongest Vertex AI fraud detection pipeline [5].

8. CONCLUSION

This paper presented an AI-driven, agent-based framework for real-time employee benefits administration using managed container services. Comprehensive evaluation across 72 enterprise implementations, four organizational scales (500 to 50,000 employees), and four enrollment scenario types demonstrates consistent 68-79% cost reductions, enrollment determination accuracy exceeding 97.9%, zero HIPAA/ERISA violations, and sub-three-minute real-time processing compared to next-day batch cycles in traditional architectures [1][6][7][12].

The quantitative cost-optimization framework, AI agent selection matrix, and platform comparison findings provide actionable guidance for benefits technology architects navigating cloud platform selection. Future research directions include adaptive pre-warming strategies for small-employer deployments, transfer learning for new-hire fraud detection, and federated ML approaches enabling cross-employer model improvement while preserving employee data privacy under HIPAA [7][9][12].

Several limitations warrant consideration. The cost model assumes a standard enrollment event profile; organisations with highly irregular enrollment spikes may experience higher cold-start penalties than reported. AI agent accuracy benchmarks were measured on English-language documents, so multinational employers with non-English dependent documentation will require language-specific model fine-tuning. Emerging state-level privacy regulations such as CCPA and VCDPA may introduce compliance requirements beyond the current framework scope. Future work should address these through multilingual document intelligence, adaptive pre-warming policies, and compliance-rule-as-code integrations [6][7].

9. ACKNOWLEDGMENTS

The author acknowledges the enterprise HR technology community for enabling access to anonymised deployment data used in constructing the DS-4 benchmark pool. Platform documentation and best-practice guidance from Amazon Web Services [3][10][15], Microsoft Azure [4][13], and Google Cloud [5] were instrumental in validating the implementation patterns described in this paper.

10. REFERENCES

[1] Society for Human Resource Management, "2024 Benefits Administration Benchmarking Survey," SHRM Research, 2024.

[2] P. Castro, V. Ishakian, V. Muthusamy, and A. Slominski, "The rise of serverless computing," Commun. ACM, vol. 62, no. 12, pp. 44-54, Dec. 2019.

[3] Amazon Web Services, "Amazon Rekognition Developer Guide," 2024. [Online]. Available: https://docs.aws.amazon.com/rekognition/

[4] Microsoft Azure, "Azure AI Document Intelligence Documentation," Microsoft Corporation, 2024. [Online]. Available: https://learn.microsoft.com/azure/ai-services/document-intelligence/

[5] Google Cloud, "Document AI Documentation," Google LLC, 2024. [Online]. Available: https://cloud.google.com/document-ai/docs

[6] U.S. Department of Labor, "ERISA Compliance Guidelines for Employee Benefits," 29 CFR Part 2520, 2024.

[7] U.S. Department of Health and Human Services, "HIPAA Privacy Rule," 45 CFR Part 164, 2024.

[8] Amazon Web Services, "Amazon DynamoDB Developer Guide," 2024. [Online]. Available: https://docs.aws.amazon.com/dynamodb/

[9] D. Sculley et al., "Hidden Technical Debt in Machine Learning Systems," in Proc. NIPS, 2015.

[10] Amazon Web Services, "Amazon SageMaker Developer Guide," 2024. [Online]. Available: https://docs.aws.amazon.com/sagemaker/

[11] E. Breck, S. Cai, E. Nielsen, M. Salib, and D. Sculley, "The ML Test Score: A Rubric for ML Production Readiness and Technical Debt Reduction," in Proc. IEEE Int'l Conf. on Big Data (BigData), pp. 1123-1132, 2017.

[12] Employee Benefit Research Institute, "2024 Employee Benefits Compliance Best Practices," EBRI Publications, 2024.

[13] Microsoft Azure, "Azure Machine Learning Documentation," 2024. [Online]. Available: https://learn.microsoft.com/azure/machine-learning/

[14] D. Bermbach, A.-S. Karakaya, and S. Buchholz, "Using application knowledge to reduce cold starts in FaaS services," in Proc. 35th ACM Symp. on Applied Computing (SAC), pp. 134-143, 2020.

[15] Amazon Web Services, "AWS ECS Best Practices Guide," 2024. [Online]. Available: https://docs.aws.amazon.com/AmazonECS/latest/bestpracticesguide/