TroveLLM Coding Agent vs. AWS Comprehend Medical

A Comparative Analysis of AI-Powered Medical Coding Solutions

Karthik Ravinutala
Venkat Timmaraju
Aaditya Muleva

January 2026

Table Of Contents

1. Executive Summary

The healthcare industry faces an unprecedented challenge in medical coding and documentation. With the global medical coding market valued at $24.83 billion in 2025 and projected to reach $42.8 billion by 2030, the demand for accurate, efficient, and cost-effective coding solutions has never been greater. Healthcare organizations spend billions annually on coding services, with claim denial costs alone reaching $262 billion per year for U.S. hospitals.

This white paper presents a comprehensive comparison between TroveLLM Coding Agent and AWS Comprehend Medical, two leading AI-powered solutions addressing this critical need. Our analysis, based on processing 202 pages across 24 clinical documents, reveals significant differences in capability, accuracy, and cost-effectiveness.

Key outcomes from the analysis include:

- *Cost Efficiency*
  TroveLLM demonstrates a 96% cost reduction compared to AWS Comprehend Medical, processing 32,000 pages for $500 versus an estimated $12,200 for AWS.
- *Code Coverage*
  TroveLLM extracts codes from 7 coding systems (LOINC, SNOMED, ICD10CM, ICD10PCS, ICD9CM, HCPCS, RxNorm) compared to AWS's 3 systems (SNOMED, ICD10CM, RxNorm).
- *Contextual Understanding*
  TroveLLM's three-stage processing pipeline provides superior contextual awareness, eliminating the need for confidence thresholds used by AWS Comprehend Medical.
- *Multilingual Support*
  TroveLLM handles English, Spanish, and French documents, including handwritten notes, while AWS is limited to English text only.
- *Processing Efficiency*
  Both solutions achieved similar processing speeds (6-7 seconds per page for TroveLLM vs. 5-6 seconds per block for AWS), but TroveLLM processed more blocks with meaningful data extraction.
- *The implications for healthcare organizations are substantial*
  Significant cost savings, improved coding accuracy, faster reimbursement cycles, and enhanced compliance with regulatory requirements.

2. Introduction: The Medical Coding Crisis

Medical coding is the backbone of healthcare revenue cycle management. Every diagnosis, procedure, and treatment must be accurately translated into standardized codes for billing, compliance, and statistical purposes. However, the industry faces critical challenges that threaten both financial stability and patient care quality.

2.1 The Scale of the Problem

The healthcare industry generates massive volumes of clinical documentation daily. U.S. hospitals alone process billions of claims annually, with each claim requiring precise coding across multiple classification systems. The complexity has grown exponentially:

- ICD-11, implemented in 2022, contains over 55,000 diagnostic codes.
- Healthcare organizations face a 30% vacancy rate for certified medical coders.
- Physicians spend an average of 90 minutes overtime daily on coding-related tasks.
- Coding errors contribute to $262 billion in annual claim denials.

2.2 The Human Cost

Beyond financial implications, manual coding processes create significant human costs:

- Clinician burnout from administrative burden.
- Delayed patient care due to documentation requirements.
- Revenue leakage from missed or incorrect codes.
- Compliance risks from coding errors.

2.3 The Technology Imperative

The convergence of artificial intelligence, natural language processing, and healthcare informatics offers a transformative solution. AI-powered medical coding systems can:

- Process clinical documents at scale with consistent accuracy.
- Extract nuanced medical information from unstructured text.
- Apply contextual understanding to code assignment.
- Operate continuously without fatigue or human error

This white paper examines two leading solutions addressing this challenge: TroveLLM Coding Agent, a purpose-built clinical AI solution, and AWS Comprehend Medical, a cloud-based NLP service from Amazon Web Services.

3. Market Analysis and Opportunity

3.1 Global Medical Coding Market

The global medical coding market represents one of the fastest-growing segments in healthcare technology. Market research indicates:

- 2024 Market Value: $24.83 billion globally.
- 2030 Projected Value: $42.8 billion.
- CAGR (2025-2030): 9.45%.

The growth is driven by several factors including the transition to ICD-11, increasing healthcare claims volume, AI adoption, and regulatory requirements for accurate documentation.
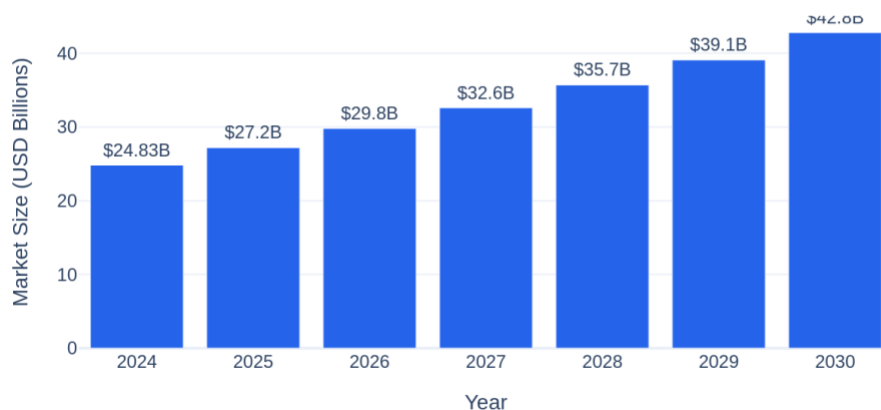

Figure 1: Global Medical Coding Market Growth Projection

3.2 Healthcare NLP Market

The Natural Language Processing (NLP) in healthcare market is experiencing even more dramatic growth:

- 2024 Market Value: $6.66 billion.
- 2034 Projected Value: $132.34 billion.
- CAGR (2025-2034): 34.74%.

This explosive growth reflects the healthcare industry's recognition that NLP technologies are essential for processing the vast amounts of unstructured clinical data generated daily.
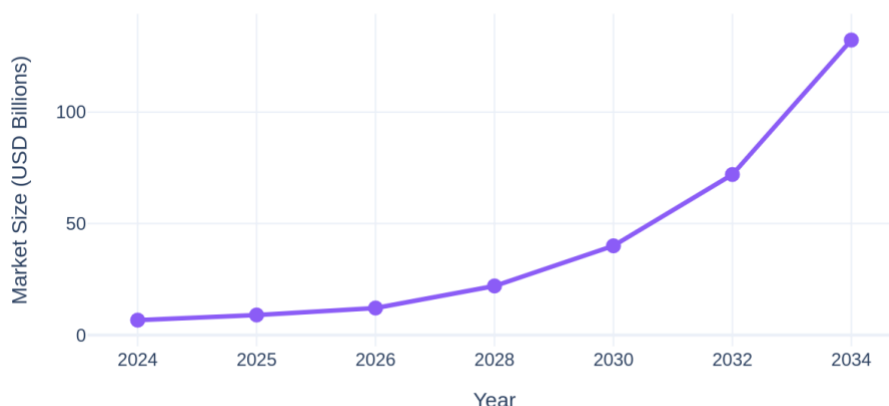

Figure 2: Healthcare NLP Market Growth Trajectory

3.3 Revenue Cycle Management Impact

Medical coding is a critical component of Revenue Cycle Management (RCM), which represents:

• 2024 RCM Market: $152.14 billion globally.
• 2034 Projected: $453.47 billion.
• CAGR: 11.54%.

Organizations investing in AI-powered coding solutions can expect significant returns through reduced administrative costs, faster claim processing, and improved reimbursement rates.
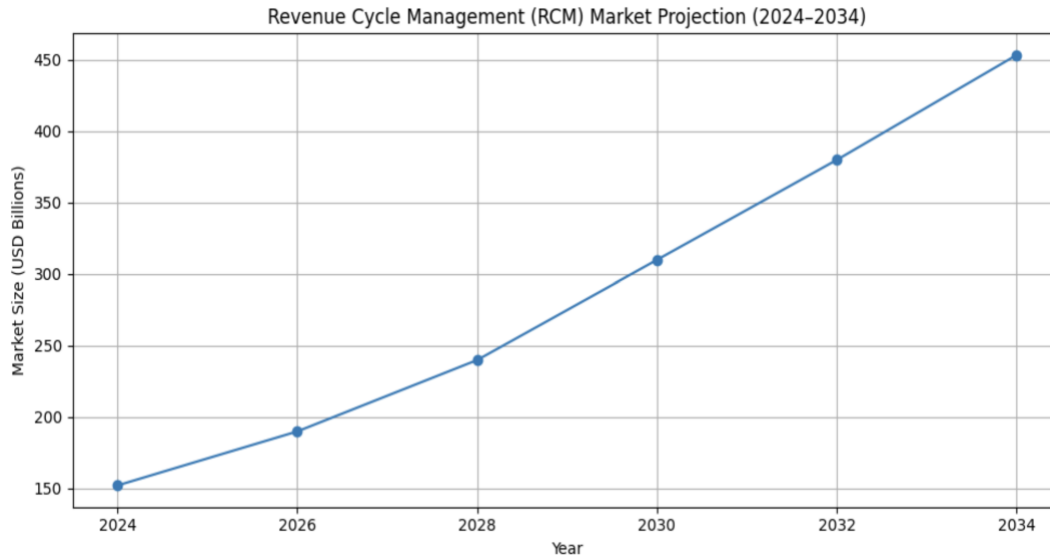
Figure 3: Revenue Cycle Management (RCM) Market Projection

4. Technology Comparison: TroveLLM vs AWS Comprehend Medical

4.1 TroveLLM Coding Agent Architecture

TroveLLM employs a sophisticated three-stage processing pipeline designed specifically for clinical documentation:

Stage 1: TrovePDF Parser
The TrovePDF parser meticulously processes clinical documents page by page, identifying and grouping text blocks in an orderly fashion. This structured approach ensures no critical information is missed during extraction.

Stage 2: Clinical Classification
Parsed blocks are sent to TroveLLM for classification into 17 clinical categories:

- Assessment and Plan
- History of Present Illness
- Medications
- Problem Lists
- Social History
- History of Procedures
- Radiology Studies
- Lab Results
- Advance Directives

- Functional Status
- Vital Signs
- Immunizations
- Pathology Studies
- Medical Equipment
- Physical Findings
- Admission Diagnosis
- Unknown/Not Functional Clinical Classification

This classification enables context-aware processing, as a single text block may contain multiple clinical classes that TroveLLM intelligently separates.

Stage 3: Code Assignment with RAG

Each classified text segment is processed through specialized Retrieval-Augmented Generation (RAG) databases for code assignment. TroveLLM maintains RAG databases for 12 coding systems:

- CVX (Vaccines)
- HCPCS (Healthcare Common Procedure Coding System)
- ICD10CM (Diagnosis codes)
- ICD10PCS (Procedure codes)
- ICD9CM (Legacy diagnosis codes)
- ICD9PCS (Legacy procedure codes)
- LOINC (Laboratory observations)
- NCI (National Cancer Institute)
- NDC (National Drug Code)
- RxNorm (Medications)
- SNOMED (Clinical terminology)
- CPT (Current Procedural Terminology)

The system returns the top 5 matches for each query, with TroveLLM's clinical intelligence selecting the most appropriate code based on full contextual understanding.

4.2 AWS Comprehend Medical Architecture

AWS Comprehend Medical is a HIPAA-eligible NLP service that extracts health data from unstructured medical text. Its architecture includes:

Named Entity Recognition (NER)

Identifies and categorizes medical entities including:

- Medical conditions
- Medications

- Dosages
- Tests, treatments, and procedures
- Protected Health Information (PHI)

Ontology Linking

Maps identified entities to standardized medical ontologies:

- ICD-10-CM (Diagnosis codes)
- RxNorm (Medications)
- SNOMED CT (Clinical terminology)

4.3 Architectural Comparison

Feature	TroveLLM	AWS Comprehend Medical
Code Systems Supported	12 systems (full coverage)	3 systems (limited)
Processing Approach	3-stage pipeline with classification	Direct entity extraction
Contextual Understanding	Full clinical context awareness	Limited context (entity-level)
Confidence Handling	Intelligent selection (no threshold)	80% threshold required
Language Support	English, Spanish, French	English only
Handwriting Support	Yes	No
Character Limit	No limit (page-based)	10,000 characters maximum

Table 1: Architectural Comparison of TroveLLM and AWS Comprehend Medical

5. Performance Benchmarking

5.1 Test Methodology

Our comparative analysis utilized a standardized dataset of clinical documents:

- Total Documents: 24 clinical documents.
- Total Pages: 202 pages.
- Document Types: Lab results, diagnosis reports, after-visit summaries.
- Average Pages per Document: 8.41.
- Text Blocks Extracted: 577 blocks.
- Average Blocks per Page: 2.85

5.2 Processing Results

TroveLLM Performance:

- Blocks Processed: 574 of 577 (99.5%)
- Blocks with No Meaningful Data: 182
- Blocks Fully Processed: 392
- Total Unique Codes Extracted: 1,811
- Total Codes (with duplicates): 1,926
- Processing Time: 6-7 seconds per page

AWS Comprehend Medical Performance:

- Blocks Processed: 573 of 577 (99.3%)
- Blocks with No Data: 147
- Blocks Below Confidence Threshold: 111
- Blocks Above Confidence Threshold: 315
- Total Unique Codes Extracted: 1,709 (at 80%+ confidence)
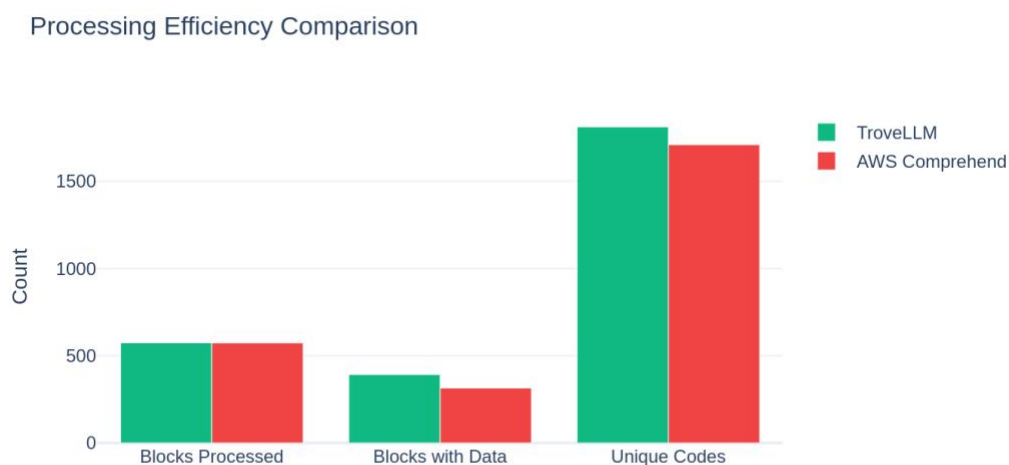- Processing Time: 5-6 seconds per text block

Processing Efficiency Comparison

Figure 4: Processing Efficiency Comparison

5.3 Code System Coverage

A significant differentiator is the breadth of coding systems each solution can extract:

TroveLLM Code Extraction by System:

- LOINC: 672 codes (laboratory observations)
- RxNorm: 520 codes (medications)
- ICD10PCS: 366 codes (procedures)
- HCPCS: 169 codes (services)
- ICD9CM: 105 codes (legacy diagnoses)
- SNOMED: 49 codes (clinical terms)
- ICD10CM: 45 codes (diagnoses)

AWS Comprehend Medical Code Extraction:

- SNOMED: 1,544 codes
- ICD10CM: 1,080 codes
- RxNorm: 149 codes
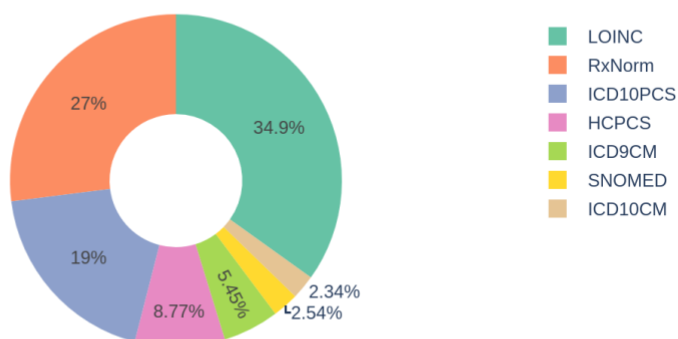
TroveLLM: Code Distribution by System

Figure 5: TroveLLM Code Distribution by System

5.4 Qualitative Observations

Manual review of 100+ AWS Comprehend Medical outputs revealed critical limitations:

1. Lack of Context
The system failed to consider surrounding context when coding entities, leading to potentially incorrect code assignments.

2. Incomplete Information Extraction
For vitals and other numeric data, AWS Comprehend Medical failed to capture the numerical details essential for accurate coding.

3. Redundant Processing
Single entities were processed multiple times with partial terms. For example, "diabetes retinal" was also searched as "retinal" and "diabetes" separately, creating inefficiency and potential for duplicate billing.

6. Cost Analysis

6.1 Pricing Models

TroveLLM Pricing

TroveLLM offers a straightforward, volume-based pricing model optimized for healthcare organizations that process large volumes of documents. Based on our testing:

- Cost for 32,000 pages: $500
- Effective cost per page: $0.0156

AWS Comprehend Medical Pricing:

AWS uses a per-unit model where 1 unit = 100 characters:

- First 1,000,000 units: $0.01 per unit
- Next 1,000,000 units: $0.005 per unit
- Over 2,000,000 units: $0.001 per unit

Based on our document characteristics (average 1,700 characters per page), the estimated cost:

- Cost for 32,000 pages: ~$12,200
- Effective cost per page: $0.38

6.2 Cost Comparison

Figure 6: Total Cost Comparison for 32,000 Pages

6.3 Return on Investment Analysis

For a typical healthcare organization processing 100,000 pages monthly:

With TroveLLM

- Monthly cost: ~$1,562
- Annual cost: ~$18,750

With AWS Comprehend Medical

- Monthly cost: ~$38,000
- Annual cost: ~$456,000

Annual Savings with TroveLLM: $437,250 (96% reduction)

6.4 Total Cost of Ownership Considerations

Beyond direct processing costs, organizations should consider:

Integration Costs
Both solutions require EHR integration, but TroveLLM's broader code system support may reduce additional integration requirements.

Training Costs
TroveLLM's intelligent selection eliminates the need for threshold tuning and manual review of low-confidence results.

Error Correction Costs
AWS Comprehend Medical's context limitations may require additional human review and correction, adding to operational costs.

Multilingual Processing
Organizations serving diverse populations benefit from TroveLLM's Spanish and French support without additional licensing.

Cost Per Page Comparison

Figure 7: Cost Per Page Comparison at Scale

7. Use Cases and Benefits

7.1 Primary Use Cases

Hospital Systems and Health Networks:

- High-volume inpatient and outpatient coding
- Multi-facility standardization
- Complex case documentation

Physician Practices and Clinics:

- Office visit coding automation
- Specialty-specific code optimization
- Same-day claim submission

Revenue Cycle Management Companies:

- Third-party coding services at scale
- Quality assurance and auditing
- Client portfolio management

Health Insurance Payers:

- Claims verification and fraud detection
- Risk adjustment validation
- Value-based care analytics

Research Institutions:

- Clinical trial documentation
- Real-world evidence generation
- Population health studies

7.2 Quantifiable Benefits

Financial Benefits:

- 96% cost reduction compared to AWS Comprehend Medical.
- Reduced claim denial rates through accurate coding.
- Faster reimbursement cycles (reduced days in A/R).
- Improved charge capture and revenue recovery.

Operational Benefits:

- 60-90% reduction in manual coding workload.
- Consistent coding quality 24/7.
- Elimination of coder shortage impact.
- Reduced overtime for clinical staff.

Compliance Benefits:

- Standardized coding practices.
- Complete audit trails.
- Reduced regulatory risk.
- Support for value-based care reporting.

Clinical Benefits:

- More accurate patient records.
- Better data for clinical decision support.
- Enhanced quality reporting capabilities.
- Improved care coordination.

7.3 Unique TroveLLM Advantages

Multilingual Processing:
Healthcare organizations serving diverse communities can process documents in English, Spanish, and French without additional solutions or manual translation.

Handwritten Note Recognition:
TroveLLM's ability to process handwritten clinical notes addresses a significant gap in healthcare documentation, where many critical notes are still handwritten.

Comprehensive Code Coverage:
With support for 12 coding systems versus AWS's 3, TroveLLM provides complete coding capability across all reimbursement and compliance requirements.

Contextual Intelligence:
The three-stage processing pipeline ensures codes are assigned based on full clinical context, not isolated entity recognition.

8. Key Buyers and Market Segments

8.1 Healthcare Provider Organizations

Hospitals and Health Systems:

- 6,100+ hospitals in the United States.
- Average coding staff: 15-50 FTEs for large systems.
- Annual coding spend: $2-10 million on major health systems.
- Decision makers: CFO, CIO, VP Revenue Cycle.

Physician Group Practices:

- 250,000+ physician practices in the U.S.
- Growing consolidation into larger groups.
- Seeking automation to address coder shortages.
- Decision makers: Practice Administrator, Office Manager.

Ambulatory Surgery Centers:

- 6,000+ ASCs nationally.
- High-volume, procedure-focused coding needs.
- Strong ROI potential from automation.
- Decision makers: Administrator, Medical Director.

8.2 Revenue Cycle Management Companies

Third-Party Coding Services:

- $10+ billion market segment.
- Scale-dependent economics favor automation.
- Competitive differentiation through technology.
- Decision makers: CEO, COO, VP Operations.

Major RCM providers include:

- Optum360 (UnitedHealth Group)
- R1 RCM
- Conifer Health Solutions
- Change Healthcare
- Ensemble Health Partners

8.3 Health Insurance Organizations

Health Plans and Payers:

- Claims verification and audit.
- Risk adjustment validation.
- Fraud detection applications.
- Decision makers: VP Claims, Chief Medical Officer.

8.4 Technology Integrators

EHR Vendors:

- Epic, Cerner, Meditech, Allscripts.
- Integration partnerships for embedded coding.
- Value-added features for customers.

Healthcare IT Companies:

- Systems integrators and consultants.
- Reseller and implementation partnerships.
- Custom solution development.

8.5 Current Market Spending

Organizations currently invest significantly in medical coding:

- Average hospital coding department: $1.5-3 million annually.
- Outsourced coding rates: $15-40 per chart (varies by complexity).
- AI coding solutions: $100,000-500,000+ implementation costs.
- Ongoing subscription fees: $0.10-1.00 per encounter.

The 96% cost advantage demonstrated by TroveLLM positions it as a compelling alternative to existing solutions and a transformative option compared to traditional manual coding operations.

9. Business Conclusion and Recommendations

9.1 Summary of Findings

Our comprehensive analysis of TroveLLM Coding Agent versus AWS Comprehend Medical reveals clear differentiation across all evaluation criteria:

Technical Superiority:
TroveLLM's three-stage processing pipeline, combined with support for 12 coding systems and multilingual capabilities, provides significantly broader and more accurate coding coverage than AWS Comprehend Medical's limited three-system approach.

Cost Leadership:
With a 96% cost reduction at scale (32,000 pages: $500 vs $12,200), TroveLLM offers an unprecedented value proposition for healthcare organizations of all sizes.

Contextual Intelligence:
TroveLLM's clinical classification and RAG-based code selection eliminates the context limitations observed in AWS Comprehend Medical, reducing coding errors and the need for manual review.

Future-Ready Capabilities:
Support for handwritten notes, multilingual processing, and comprehensive code system coverage positions TroveLLM for evolving healthcare documentation needs.

9.2 Strategic Recommendations

For Healthcare Provider Organizations:

- Evaluate current coding costs and error rates to establish baseline.
- Pilot TroveLLM on representative document samples.
- Calculate specific ROI based on volume and complexity.
- Develop implementation roadmap with EHR integration.

For RCM Companies:

- Assess competitive advantage from 96% cost reduction.
- Explore white-label or partnership opportunities.
- Quantify capacity increase potential with automation.
- Plan service offering expansion with multilingual support.

For Health Plans:

- Evaluate claims verification use cases.
- Assess risk adjustment validation applications.
- Consider fraud detection capabilities.
- Explore population health analytics potential

9.3 The Path Forward

The medical coding market is at an inflection point. With a $24.83 billion market growing at nearly 10% annually, and acute coder shortages driving operational challenges, AI-powered solutions are no longer optional—they are essential.

TroveLLM Coding Agent represents a new paradigm in clinical coding technology: purpose-built for healthcare, optimized for accuracy, and priced for accessibility. Organizations that embrace this technology will gain significant competitive advantages through reduced costs, improved accuracy, faster reimbursement, and enhanced compliance.

The data is clear: TroveLLM delivers superior performance at a fraction of the cost. The question is not whether to adopt AI-powered medical coding, but how quickly organizations can implement it to capture these benefits.

10. References

1. Grand View Research. (2025). Medical Coding Market Size, Share & Trends Analysis Report. https://www.grandviewresearch.com/industry-analysis/medical-coding-market

2. Mordor Intelligence. (2025). Medical Coding Market - Growth, Trends, COVID-19 Impact, and Forecasts. https://www.mordorintelligence.com/industry-reports/medical-coding-market

3. Precedence Research. (2025). Medical Coding Market Size, Share, Growth, Report 2024-2033. https://www.precedenceresearch.com/medical-coding-market

4. Markets and Markets. (2025). Healthcare NLP Market - Global Forecast to 2030. https://www.marketsandmarkets.com/Market-Reports/healthcare-lifesciences-nlp-market.html

5. Towards Healthcare. (2025). Natural Language Processing in Healthcare and Life Sciences Market. https://www.towardshealthcare.com/insights/natural-language-processing-nlp-in-healthcare-and-life-sciences-market-sizing

6. AWS. (2025). Amazon Comprehend Medical Pricing. https://aws.amazon.com/comprehend/medical/pricing/

7. Invensis. (2025). Coding and Revenue Cycle Management in Healthcare. https://www.invensis.net/blog/coding-and-revenue-cycle-management

8. Virtue Market Research. (2025). Global Autonomous Medical Coding Market. https://virtuemarketresearch.com/report/global-autonomous-medical-coding-market

9. Coherent Market Insights. (2025). Computer-Assisted Coding Market Analysis. https://www.coherentmarketinsights.com/industry-reports/medical-coding-software-market

10. PMC. (2024). Revenue Cycle Management in Healthcare Organizations. https://pmc.ncbi.nlm.nih.gov/articles/PMC11219169/

11. AAPC. (2025). Medical Coding Industry Standards and Best Practices. https://www.aapc.com/

12. CMS. (2025). Healthcare Common Procedure Coding System. https://www.cms.gov/medicare/coding-billing/healthcare-common-procedure-system