Journal of Recent Trends in Computer Science and Engineering (JRTCSE)
Vol. 14, No. 1, January-February, 2026, PP. 7-15
ISSN: 2322-0872
https://jrtcse.com
DOI: https://doi.org/10.70589/JRTCSE.2026.14.1.2

# Domain-Specific Language Models in Healthcare: A Framework for Trust, Governance, and Safe Deployment

**Preeti Tupsakhare,**
Engineer Lead - Medical Benefit Management Information Technology,
Elevance Health, USA

**Abstract**

General-purpose large language models (LLMs) have demonstrated significant potential in healthcare applications but remain ill-suited for high-stakes clinical environments due to limited domain awareness, lack of explainability, and opaque reasoning. This white paper introduces Domain-Specific Language Models (DSLMs) as a purpose-built alternative designed to meet the trust, safety, and regulatory demands of healthcare systems. It presents a comprehensive framework for the development, validation, and deployment of DSLMs, emphasizing data governance, model transparency, bias mitigation, and regulatory alignment. The paper outlines practical strategies for integrating DSLMs across clinical and administrative workflows while maintaining human oversight and patient safety. Collectively, this framework provides healthcare organizations with a scalable and responsible pathway for adopting AI in mission-critical settings.

**Keywords:** Domain-specific Language models, AI, Artificial Intelligence, customized healthcare LMs

**Citation:** Preeti Tupsakhare. (2026). Domain-Specific Language Models in Healthcare: A Framework for Trust, Governance, and Safe Deployment. Journal of Recent Trends in Computer Science and Engineering, 14(1), 7-15.

## 1.Introduction

The rapid evolution of generative AI has revolutionized healthcare data interpretation, medical research, and patient engagement. However, the general-purpose nature of most large language models introduces significant risks when applied to sensitive, high-stakes healthcare environments. Challenges such as hallucinated clinical insights, biased or non-representative outputs, limited explainability, and misalignment with healthcare regulations can compromise patient safety, erode clinician trust, and expose organizations to regulatory and ethical risk.

Domain-Specific Language Models (DSLMs) offer a targeted solution—trained on validated medical corpora, aligned with clinical workflows, and designed to operate within established ethical and regulatory boundaries. Their emergence represents a paradigm shift from broad AI capability to trustworthy, domain-grounded intelligence tailored for healthcare decision support.

This white paper addresses the practical challenges of adopting DSLMs in real-world healthcare settings, including data governance, model transparency, bias mitigation, validation, and regulatory compliance. It proposes a structured, phase-wise implementation framework that enables organizations to safely pilot, evaluate, and scale DSLMs—beginning with low-risk administrative use cases and progressing toward clinician-facing applications under human oversight. By outlining clear guardrails and deployment strategies, the paper provides a roadmap for responsible, scalable adoption of DSLMs across healthcare systems.

## 2. Understanding Domain Specific Language Models (DSLMs)

Domain-Specific Language Models (DSLMs) differ from general-purpose LLMs by being intentionally designed for the linguistic, contextual, and regulatory requirements of healthcare. Rather than optimizing for broad language capability, DSLMs prioritize clinical accuracy, traceability, and safety [1].

Key characteristics of DSLMs include:

- **Specialized Training Data:** Curated datasets derived from verified medical literature, clinical notes, and anonymized electronic health records (EHRs) reduce reliance on non-authoritative sources and limit spurious content generation [2].
- **Controlled Vocabulary and Ontologies:** Integration with medical ontologies such as UMLS, SNOMED CT, and ICD-10 ensures consistent terminology usage and improves contextual and semantic accuracy.
- **Enhanced Validation Pipelines:** Continuous benchmarking against domain-specific evaluation datasets and real-world clinical cases improves factual reliability and detects performance drift.
- **Hallucination Mitigation Mechanisms:** Constrained domain training, retrieval-augmented generation, and output verification steps reduce the likelihood of clinically unsafe hallucinations compared to general-purpose models.
- **Domain-Aware Guardrails:** Predefined ethical, clinical, and contextual constraints limit unsupported recommendations and reinforce use as clinical decision support under human oversight [2].

Representative DSLMs, including Med-PaLM 2, BioGPT, and ClinicalBERT, DrugGPT have demonstrated improved performance over general-purpose LLMs in medical question answering, clinical document summarization, and coding support tasks. These characteristics collectively position DSLMs as a more reliable foundation for AI deployment in healthcare environments where errors and hallucinations can have material clinical consequences [1].

## 3. Examples Of Domain Specific DSLMs Across the Industries

As organizations confront the limitations of general-purpose LLMs, several industries have adopted domain-specific language models tailored to regulatory, operational, and domain constraints. Table 1 presents illustrative examples of such models, providing context for how domain specialization can inform the development of governed DSLMs in healthcare.

**Table: 1**

| Model | Domain | Provider/Notes | Highlights / Suitability |
|---|---|---|---|
| **BloombergGPT** | Finance / Banking | Developed by Bloomberg L.P.; trained on ~363B tokens of proprietary financial data combined with a general corpus [7] | Demonstrates strong financial-domain performance<br><br>(market analysis, sentiment, regulatory filings).<br><br>Serves as a mature example of a vertical model,<br><br>though not tailored for healthcare or insurance workflows [7]. |
| Med-PaLM 2 | Healthcare / Medical | Developed by Google Research / DeepMind; fine-tuned for medical question answering | Well-suited for clinical summarization and medical Q&A. Relevant for care management and clinician support, but less aligned with payer-specific workflows such as claims or utilization management. |
| Palmyra Med 70B | Healthcare | Referenced by NVIDIA for deep medical reasoning use cases | Strong medical domain depth (genomics, clinical trials, biomedical literature). May require additional adaptation for payer and administrative healthcare contexts. |
| Lexis+ AI | Legal / Compliance | Developed by LexisNexis; trained on proprietary legal and regulatory content [8] | Strong example of a vertical model for regulated domains. Relevant to healthcare for contract analysis, regulatory interpretation, and provider-network legal workflows [8]. |
| DrugGPT | BioMedical | Open-source model focused on drug | Narrow but deep biomedical focus. Less immediate applicability for payer workflows, but demonstrates the |

| | | | |
|---|---|---|---|
| | | discovery and molecular research [9] | effectiveness of highly specialized domain models [9]. |
| FinGPT | Finance / FinTech | Open-source, community-driven fine-tuned models [10] | Illustrates flexibility of open-domain specialization. Customizable but not turnkey for healthcare or insurance without significant governance and domain adaptation |

## 4. Core Challenges in Healthcare Dslm Development

Building Domain-Specific Language Models (DSLMs) for healthcare introduces challenges distinct from general AI deployment:

- **Data Privacy & Compliance:** Ensuring strict adherence to HIPAA and other regulatory requirements, including robust de-identification, secure data handling, and patient consent tracking [3].
- **Bias Mitigation:** Addressing imbalances and historical disparities in source datasets that could perpetuate inequities in clinical insights or care recommendations [3].
- **Explainability & Accountability:** Providing transparent model reasoning, confidence indicators, and audit trails necessary to support clinician trust and regulatory scrutiny [6].
- **Clinical Validation & Safety Oversight:** Demonstrating clinical reliability through rigorous benchmarking, human-in-the-loop evaluation, and clear safeguards to prevent unsafe or unsupported recommendations [6].
- **Integration Complexity:** Aligning DSLM outputs with existing EMR/EHR systems, payer platforms, and clinical workflows without increasing operational burden.
- **Resource Intensity:** Balancing model performance, latency, and computational cost to ensure scalability in real-world healthcare environments [6].

## 5. Governance Framework For Trustworthy Healthcare DSLMs

A trustworthy DSLM ecosystem requires an end-to-end governance framework structured around six core pillars:

### 5.1. Data Governance

- Data provenance and lineage tracking across all training and inference datasets [2]
- Strict anonymization, de-identification, and differential privacy standards [4]
- Version control and documentation of training data sources [4]

### 5.2. Model Governance

- End-to-end lifecycle management, including training, validation, deployment, and retraining [2]
- Continuous model drift detection and performance monitoring
- Role-based access control for model updates and production releases

### 5.3. Clinical Governance

- Clear delineation of clinical ownership and accountability for DSLM-supported decisions
- Human-in-the-loop review for safety-critical use cases
- Alignment with clinical guidelines, pathways, and standards of care

### 5.4. Ethical Governance

- Independent ethics review boards for AI use cases
- Regular algorithmic fairness and bias audits
- Inclusion of underrepresented populations in training and evaluation datasets

### 5.5. Operational Governance

- Real-time system monitoring, alerting, and rollback mechanisms
- Comprehensive audit logs supporting human oversight and post-incident review [4]
- Defined escalation and incident response pathways for model anomalies

### 5.6. Regulatory Governance

- Alignment with applicable regulations and guidance, including the FDA AI/ML-Based SaMD Action Plan, EU AI Act, and WHO Ethics Guidance
- Documentation to support regulatory submissions, audits, and AI incident reporting

## 6. Implementation Roadmap

To balance innovation with patient safety and regulatory accountability, this white paper proposes a phased implementation framework that enables progressive adoption of Domain-Specific Language Models in healthcare environments [4] [5].

- **Phase 1: Data Curation & Governance Foundation**

  Establish high-quality, labeled datasets derived from validated medical sources, supported by strong governance controls including consent management, data provenance, de-identification, and privacy safeguards [4][5].

- **Phase 2: Model Training & Clinical Validation**

  Apply transfer learning from vetted foundational or open-source models, followed by domain-specific fine-tuning and rigorous validation against benchmark datasets and representative real-world clinical scenarios [2] [5].

- **Phase 3: Governance & Safety Layer Integration**

  Embed cross-cutting governance capabilities, including explainability mechanisms, bias and drift monitoring, audit logging, access controls, and defined accountability structures [3].

- **Phase 4: Controlled Deployment & Risk-Based Scaling**

  Introduce DSLMs through controlled pilots in low-risk administrative and operational workflows (e.g., documentation summarization, medical coding), with human-in-the-loop oversight prior to expansion into clinician-facing use cases [5][6].

- **Phase 5: Continuous Oversight & Lifecycle Management**

  Maintain ongoing performance monitoring, clinician and user feedback loops, drift detection, and periodic revalidation to ensure sustained accuracy, safety, and regulatory compliance.

## 7. Example Use Cases

### 7.1. Clinical & Care Management

- Summarize longitudinal patient histories and care gaps for nurse case managers and care coordinators.
- Generate personalized outreach templates for chronic disease and population health management programs.
- Suggest care plan adjustments or next-best actions based on utilization patterns, prior authorization history, and clinical guidelines, with human review.

### 7.2. Claims & Utilization Management

- Automatically classify and triage claims with a high likelihood of denial or rework using unstructured claim narratives.
- Support fraud, waste, and abuse (FWA) detection by extracting patterns from free-text provider notes and claim descriptions.
- Draft appeal letters or utilization review summaries for denied claims, citing relevant policies and supporting evidence.

### 7.3. Member & Provider Experience

- Power natural language assistants that understand payer–provider context, benefits structures, and coverage rules.
- Enable secure, multi-turn conversations that retrieve and summarize members or provider data to resolve inquiries efficiently.
- Deliver personalized, context-aware member education, such as benefit explanations, preventive care reminders, and care navigation guidance.

### 7.4. Multimodal Clinical Insights

- Combine text-based DSLMs with imaging reports, lab trends, and remote monitoring data to support care coordination and risk stratification.

### 7.5. Federated Intelligence Across Organizations

- Enable privacy-preserving learning across providers, payers, and care networks to improve utilization insights and population health modeling without centralizing sensitive data.

Across all usecases, healthcare DSLMs must operate within strict security and ethical boundaries:

- Adhere to HIPAA, SOC 2, and organizational security standards for protected health information (PHI).
- Minimize hallucination risk by grounding responses in verified internal data sources and clinical guidelines.
- Provide explainability and traceability, particularly for clinical or utilization-related outputs.
- Conduct regular bias audits to ensure equitable treatment across diverse members and provider populations [4].

## 8. Conclusion

The growing adoption of generative AI in healthcare has highlighted both its transformative potential and its inherent risks. General-purpose language models, while powerful, lack the domain grounding, transparency, and governance required for high-stakes healthcare environments. Domain-Specific Language Models represent a necessary evolution—designed to operate within clinical context, regulatory constraints, and ethical boundaries.

This white paper has presented a comprehensive framework for the responsible development and deployment of DSLMs, spanning data stewardship, multi-layer governance, rigorous validation, and phased implementation. When applied thoughtfully, DSLMs can deliver tangible value across healthcare operations, including care management support through patient history summarization and outreach generation, claims and utilization optimization through denial prediction and appeal drafting, and improved member and provider experiences via context-aware conversational assistants.

Looking forward, advances in multimodal intelligence and federated learning will further expand these capabilities—enabling DSLMs to reason across clinical text, imaging, utilization patterns, and real-world signals while preserving patient privacy. However, the long-term success of these systems will depend not solely on technical innovation, but on disciplined governance, continuous oversight, and human-centered design.

By embedding security, explainability, bias mitigation, and clinical accountability into every stage of the model lifecycle, healthcare organizations can operationalize DSLMs as trusted copilots rather than autonomous decision-makers. In doing so, DSLMs can help reduce administrative burden, support more informed decision-making, and contribute to safer, more equitable, and more efficient healthcare delivery.

## References

Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. 2021. Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing. ACM Trans. Comput. Healthcare 3, 1, Article 2 (January 2022), 23 pages. https://doi.org/10.1145/3458754.

Benyou Wang, Qianqian Xie, Jiahuan Pei, Zhihong Chen, Prayag Tiwari, Zhao Li, and Jie Fu. 2023. Pre-trained Language Models in Biomedical Domain: A Systematic Survey. ACM Comput. Surv. 56, 3, Article 55 (March 2024), 52 pages. https://doi.org/10.1145/3611651

M. Firdaus and K. -H. Rhee, "Towards Trustworthy Collaborative Healthcare Data Sharing," *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Istanbul, Turkiye, 2023, pp. 4059-4064, doi: 10.1109/BIBM58861.2023.10385319.

Yacine Jernite, Huu Nguyen, Stella Biderman, Anna Rogers, Maraim Masoud, Valentin Danchev, Samson Tan, Alexandra Sasha Luccioni, Nishant Subramani, Isaac Johnson, Gerard Dupont, Jesse Dodge, Kyle Lo, Zeerak Talat, Dragomir Radev, Aaron Gokaslan, Somaieh Nikpoor, Peter Henderson, Rishi Bommasani, and Margaret Mitchell. 2022. Data Governance in the Age of Large-Scale Data-Driven Language Technology. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). Association for Computing Machinery, New York, NY, USA, 2206–2222. https://doi.org/10.1145/3531146.3534637

Kosar, T., Martı, P.E., Barrientos, P.A. and Mernik, M., 2008. A preliminary study on various implementation approaches of domain-specific language. *Information and software technology*, *50*(5), pp.390-405.

Nazi ZA, Peng W. Large Language Models in Healthcare and Medical Domain: A Review. *Informatics*. 2024; 11(3):57. https://doi.org/10.3390/informatics11030057

A. Wagh and A. Mavle, "Review of Language Models in the Financial Domain," *2024 Intelligent Systems and Machine Learning Conference (ISML)*, Hyderabad, India, 2024, pp. 140-145, doi: 10.1109/ISML60050.2024.11007354.

Bloom, B. (2023). How Technology Can Empower Diversity in Law Firms. *Solic. J.*, *166*, 62.

Li, Y., Gao, C., Song, X., Wang, X., Xu, Y., & Han, S. (2023). Druggpt: A gpt-based strategy for designing potential ligands targeting specific proteins. *bioRxiv*, 2023-06.

Luukkonen, R., Komulainen, V., Luoma, J., Eskelinen, A., Kanerva, J., Kupari, H. M., ... & Pyysalo, S. (2023, December). FinGPT: Large generative models for a small language. In *Proceedings of the 2023 conference on empirical methods in natural language processing* (pp. 2710-2726).