# Position: Biosafety-Critical Adjacent Technologies are Critical for Scalable and Safe Clinical Multi-modal LLM Deployment

**Azmine Toushik Wasi*** **and Md. Iqramul Hoque**
[1]Computational Intelligence and Operations Laboratory (CIOL)
[2]Shahjalal University of Science and Technology
Correspondence to: azmine32@student.sust.edu

## Abstract

Rise of Multimodal Large Language Models (MLLMs) marks a paradigm shift in healthcare, with the potential to revolutionize diagnostics, personalized medicine, and predictive analytics. Yet, the transformative power of MLLMs cannot be realized in isolation. In this position paper, we argue that clinical impact of AI hinges not only on the models themselves but on an integrated ecosystem of enabling technologies that ensures bio-safety and governance. These include high-fidelity data curation pipelines, multimodal data lakes, model monitoring and audit tools, secure API infrastructures, workflow orchestration layers, and seamless connectors to Electronic Health Record and Picture Archiving and Communication System platforms. Far from being peripheral, these adjacent technologies are forming a critical foundation for scalable, safe, and trustworthy clinical deployment. As this ecosystem rapidly matures into a distinct sector within digital health, strategic investment and cross-disciplinary collaboration will be essential for healthcare systems and technology vendors aiming to harness the full value of MLLMs in real-world settings.

## 1 Introduction

Multimodal Large Language Models (MLLMs) represent a pivotal advancement in artificial intelligence, particularly within clinical domains, as they combine the reasoning strengths of traditional LLMs with the capacity to process diverse modalities [26, 65]. These models can ingest and interpret heterogeneous inputs, including clinical notes, research literature, medical images (X-rays, MRIs, pathology slides), time-series data from wearables or EHRs, audio (heart/lung sounds, patient interviews), and video from neurological or surgical contexts [47, 67]. Owing to this versatility, MLLMs are increasingly regarded as Foundation Models that underpin a broad spectrum of healthcare applications [47, 67]. Their clinical potential is wide-ranging [46], enabling integrated diagnostics, such as differentiating asthma, COPD, and pneumonia by combining history, imaging, and biometrics [53], and supporting personalized oncology treatments informed by genomic, dietary, and lifestyle factors [53, 22]. MLLMs also advance predictive healthcare by enabling early disease detection, patient deterioration forecasting, and outbreak monitoring through multimodal fusion of EHRs, meteorological, and sentiment data [53]. Emerging systems such as *XMedGPT* exemplify this paradigm, attaining state-of-the-art performance in high-stakes tasks like cancer recurrence prediction while improving interpretability with multimodal explanations [66].

Despite their impressive capabilities, real-world deployment of MLLMs in clinical practice remains fraught with challenges, as general-purpose models often lack the transparency, domain customization, and explainability required in high-risk healthcare [40]. Their inherent opacity, functioning

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: NeurIPS 2025 Workshop on Biosecurity Safeguards for Generative AI.


Figure 1: **Emerging Ecosystem of Clinical MLLMs with Biosafety-Critical Adjacent Technologies**: Pillars, Current Situation, and Recommendations.

as *black boxes*, undermines clinician trust and regulatory acceptance [9, 55], since without clear reasoning, professionals hesitate to rely on outputs for diagnosis or treatment despite high reported accuracy [9]. This trust gap is further compounded by accountability concerns, as opaque errors obscure responsibility and raise significant ethical and legal issues [4]. Consequently, innovation and investment are shifting from model development alone toward workflow integration supported by robust, interoperable infrastructures [34]. Such infrastructures require reliable data acquisition and curation, secure and scalable deployment environments, continuous monitoring, and regulatory alignment [51]. We argue that enabling technologies, multimodal data lakes, workflow middleware, secure API gateways, monitoring tools, and connectors to EHR and PACS systems, constitute a new industry segment, the *Clinical MLLM Adjacent*. This sector is rapidly emerging, driven by strategic investments from major healthcare players such as Bayer, Medtronic, and AstraZeneca, who are embedding AI across clinical and operational infrastructures [68, 36, 17].

In this position paper, we argue that as MLLMs become increasingly commodified with the rise of accessible models such as GPT-4o and Gemini [12], the locus of clinical innovation will shift toward the surrounding ecosystem. Building resilient, interoperable, secure, and biosafety-aligned infrastructures for AI deployment will be the key differentiator determining which organizations can scale these technologies effectively [21]. We reframe the discussion on MLLMs in clinical settings by emphasizing the strategic role of adjacent technologies (Figure 1). Future success in healthcare AI will hinge on proactive investment, human-centered design, cross-sector collaboration, and regulatory alignment. By delineating this emerging ecosystem, we outline a roadmap for stakeholders to enable safe, equitable, and scalable clinical MLLM deployment.

## 2 Background: Pillars of the Clinical MLLM Ecosystem

Successful deployment of MLLMs in healthcare is enabled by a synergistic suite of adjacent technologies (Table 3), which collectively form the decisive infrastructure for delivering accurate, reliable, and ethically sound clinical outcomes. A comprehensive analysis of these emerging enablers is presented in Appendix B.

■ **Specialized Data Curation Tools.** Data curation is a foundational pillar for clinical MLLMs, converting heterogeneous sources, EHRs, imaging, literature, and trial data, into high-quality, reproducible datasets through de-identification, expert annotation, and advanced techniques such as instruction augmentation and chain-of-thought labeling[28, 62, 63, 41]. Well-curated datasets reduce hallucinations, enhance factual accuracy, support clinical plausibility, and mitigate algorithmic bias, yet curation remains resource-intensive, with annotation bottlenecks, privacy-utility trade-offs, and the risk of obsolescence as standards evolve[52, 69, 53, 51]. Industrial platforms such as Shaip and

Elucidata's Polly exemplify scalable, compliant pipelines for multimodal data ingestion, annotation, and harmonization, collectively forming evolving infrastructures that are essential enablers of MLLM accuracy, trust, and long-term clinical utility[62, 19].

■ **Multimodal Data Lakes.**Multimodal data lakes provide the essential infrastructure for clinical MLLMs by unifying heterogeneous datasets, including EHRs, imaging, omics, and clinical notes, within secure, scalable repositories while enabling parallel processing and regulatory compliance via HITRUST and HIPAA standards[19]. Traditional warehouses or siloed systems are insufficient for the cross-modal demands of MLLMs, making purpose-built data lakes critical for harmonizing disparate data points into clinically meaningful representations[45]. Leading platforms exemplify practical implementation: AWS for Health offers HealthOmics and HealthImaging alongside partner solutions and HealthLake services that convert legacy records into FHIR-compliant formats[61, 60], while Snowflake's AI Data Cloud integrates unstructured and structured data for real-time insights and secure collaboration[30]. Notwithstanding these advances, ongoing challenges in interoperability, standardization, and governance remain, underscoring the centrality of multimodal data lakes as the backbone for safe, scalable, and clinically actionable MLLM deployment.

■ **Robust Model Monitoring Platforms.**Robust model monitoring platforms are critical for the safe deployment of clinical MLLMs, providing continuous oversight of performance, reliability, and safety, including detection of data drift, distribution shifts, model degradation, and anomalies while delivering actionable insights for rapid issue resolution[3]. Monitoring is particularly essential for managing hallucinations, where MLLMs generate medically implausible outputs, and for mitigating bias from imbalanced training data to ensure equitable model behavior[69, 52]. Explainable AI (XAI) underpins these efforts by making *black box* decisions interpretable through techniques like SHAP, saliency maps, and example-based explanations, fostering clinician trust, safety, and regulatory compliance[9, 1]. Leading solutions, including Fiddler AI, Evidently AI, and Cognome's ExplainerAI™, integrate monitoring, drift detection, bias assessment, and compliance features, collectively forming an indispensable safeguard for clinical accuracy, patient safety, and adherence to evolving ethical and regulatory standards[3, 2, 15].

■ **Secure API Gateways.**Secure API gateways are essential for MLLM deployment in healthcare, providing controlled, scalable access to AI services while safeguarding sensitive patient data through zero-trust security, authentication, and authorization mechanisms[14, 63]. These gateways optimize performance and cost via traffic management, policy enforcement, and semantic caching, yet their configuration for heterogeneous workloads and diverse modalities presents operational and governance challenges[13]. Functioning as critical *trust boundaries*, they enforce encryption, role-based access, de-identification of PHI, and comprehensive logging to maintain compliance and mitigate risks[37, 13]. Leading platforms such as Google Cloud Healthcare API, Apigee, and KrakenD illustrate how robust security frameworks combined with AI-specific management enable reliable, auditable, and ethically aligned MLLM integration[14, 37].

■ **Workflow Integration Middleware.**Workflow integration middleware is essential for clinical MLLM adoption, enabling secure, seamless data exchange between EHRs, PACS, medical devices, and cloud-based systems while supporting incremental IT modernization[50]. It facilitates AI-driven automation of repetitive tasks such as scheduling, billing, and documentation, thereby reducing administrative burden, improving resource utilization, and mitigating clinician burnout[64, 49]. Beyond integration, middleware orchestrates cohesive AI-managed workflows by coordinating tasks, monitoring performance, and routing outputs based on real-time data, allowing MLLM predictions to trigger downstream actions[8]. Leading platforms such as Core Mobile PCSIP, Orases, NextGen Mirth, and Cflow illustrate these capabilities, yet their effectiveness relies on sustained governance, interoperability, and alignment with institutional workflows[29, 50, 24].

■ **Specialized EHR/PACS Connectors**Specialized EHR/PACS connectors form the clinical data backbone for MLLMs, enabling standardized, real-time access to electronic health records and high-resolution medical imaging while integrating DICOM, HL7, and FHIR-compliant data to prevent workflow disruptions[27, 43, 56]. Notwithstanding these advantages, heterogeneity across vendor systems and evolving standards presents interoperability challenges requiring continuous updates and governance[27, 16]. Beyond integration, connectors empower AI-driven workflows by enabling MLLMs to generate structured reports, highlight critical findings, and automate administrative tasks while preserving clinician oversight[44]. Leading implementations, including Medicai, Purview, and Dataloop HL7 FHIR Model V1, demonstrate how standardized connectors facilitate bidirectional

data flow, yet their effectiveness depends on robust governance, adherence to evolving standards, and careful alignment with clinical workflows[43, 54, 18].

## 3 Emerging *Clinical MLLM Adjacent* Industry Sector

Colelctive growth and strategic significance of these technologies highlight the emergence of a distinct *Clinical MLLM Adjacent* sector with unique market dynamics and adoption challenges.

✓✓ **Market Dynamics and Investment Trends.** Healthcare AI market is experiencing substantial growth, projected to increase from USD 21.66 billion in 2025 to USD 110.61 billion by 2030, reflecting a robust CAGR of 38.6% [25]. Broader estimates place the 2024 market at approximately USD 29.2 billion, with projections exceeding USD 500 billion over the next decade [6]. This trajectory is driven by significant public and private investments, accelerated AI adoption, and advancements in human-aware AI systems [25].

Specific segments exhibit notable activity (Table 1): AI-based clinical trials solutions are valued at USD 2.88 billion in 2025, growing to USD 17.40 billion by 2034 at a CAGR of 22.13% [58]; AI in diagnostics is projected to grow from USD 1.97 billion in 2025 to USD 5.44 billion by 2030 at 22.46% CAGR [48]. AI clinical care market is estimated at USD 11.35 billion in 2025 and expected to reach USD 95.15 billion by 2034 at 26.65% CAGR [23]. The global healthcare API market is forecasted to grow from USD 1.38 billion in 2025 to USD 1.92 billion by 2033 (CAGR 4.2%) [59], with alternative estimates placing it at USD 343.8 million by 2033 (CAGR 3.7%) [57]. The healthcare middleware market is projected from USD 3.0 billion in 2023 to USD 7.06 billion by 2032 at 9.97% CAGR [31], while the PACS systems market is expected to expand from USD 5.41 billion in 2024 to USD 7.601 billion by 2033, with departmental PACS growing from USD 2.86 billion to USD 4.71 billion over the same period [32, 33].

Venture capital underscores the sector's growing importance: digital health funding hit $6.4 billion in H1 2025, with AI startups receiving 62% ($3.95 billion), raising $34.4 million per round versus $18.8 million for non-AI firms [39]. Mega deals exceeding $100 million increasingly target AI companies, reflecting investor confidence, while tech giants like Google, Microsoft, IBM, and NVIDIA invest heavily in healthcare AI models and deployment infrastructure [6]. This funding pattern highlights that value lies not only in AI models but also in supporting infrastructure, data curation, multimodal data lakes, secure API gateways, and workflow integration, which enables scalable, compliant deployment, creating a distinct market segment. Economically, integrating AI via this infrastructure enhances diagnostic accuracy, personalizes treatment, improves operational efficiency, and reduces costs by preventing complications, optimizing resource use, and automating administrative tasks, freeing clinicians for patient care [35, 49].

✓✓ **Regulatory Landscape and Governance Imperatives.** Regulatory landscape for AI in healthcare (Table 2) is rapidly evolving to accommodate adaptive technologies, with the FDA regulating AI software as a Medical Device (SaMD) and introducing Predetermined Change Control Plans (PCCPs) to streamline post-market modifications while ensuring rigorous oversight [4, 11]. Compliance with data privacy laws such as HIPAA and GDPR, alongside global ethical guidance from bodies like WHO, mandates robust protections for PHI, bias mitigation, and cybersecurity safeguards [4, 51]. Notwithstanding these requirements, comprehensive governance frameworks throughout the MLLM lifecycle enhance data quality, transparency, and explainability, reducing breaches and fostering trust in clinical AI outputs [51]. Consequently, evolving regulations do not merely impose constraints but actively shape the *clinical MLLM adjacent* ecosystem by driving demand for monitoring, secure integration, and data provenance technologies, thereby converting compliance into a strategic advantage. Yet, organizations must balance regulatory adherence with operational flexibility, as overly rigid implementations could impede innovation and model agility.

✓✓ **Challenges and Opportunities for Widespread Adoption.** Despite substantial investment, the adoption of clinical MLLMs and adjacent technologies is constrained by technical, human, and organizational challenges, including limited interoperability across EHRs and PACS, diverse data formats, high IT upgrade costs, and persistent data quality and security concerns [27, 10]. Human factors further complicate implementation, as clinician resistance, poor UX, and opaque AI outputs can hinder trust and adoption, with over 63% of AI projects reportedly failing due to these issues [34, 10]. Moreover, workforce impacts necessitate retraining to manage AI outputs effectively, while ethical considerations, such as patient privacy, consent, and liability, demand robust governance [5, 4]. Consequently, achieving *human-AI symbiosis* through human-centered design, explainable AI, and

seamless workflow integration is essential to ensure that MLLMs augment clinical care, foster trust, and catalyze sustainable cultural transformation [34, 5].

## 4 Solutions and Recommendations

To fully realize the potential of clinical MLLMs, a multi-faceted strategy emphasizing integrated infrastructure, human-centered design, collaboration, and robust governance is essential. A concise overview is presented here, while a more detailed discussion is available in Appendix C.

♦ **Strategic Investment in Integrated Infrastructure.** Organizations should prioritize investments across the entire *MLLM adjacent* ecosystem, recognizing that isolated MLLMs provide limited value without robust support. Market trends highlight an *infrastructure premium*, with AI-enabled solutions and data infrastructure attracting the majority of digital health funding [39]. Investments should span specialized data curation tools for high-quality, unbiased, de-identified datasets, multimodal data lakes for scalable clinical data access [52], and secure API gateways for compliant and efficient data flow [14]. Public-private partnerships and cloud-based solutions further enhance scalability, cost-efficiency, and computational capacity [25].

♦ **Prioritizing Human-Centered AI Design and Explainability.** All adjacent technologies interacting with clinicians must emphasize usability, trust, and explainability, as human factors often determine adoption success [34]. Transparent, intuitive interfaces integrated into existing workflows reduce disruption, while explainable AI (XAI) ensures clinicians understand MLLM outputs, supporting patient safety and regulatory compliance [9]. Early clinician engagement, silent trials, and embedding XAI within monitoring platforms help secure buy-in and map workflow pain points [10].

♦ **Cross-Stakeholder Collaboration and Standardization.** Widespread adoption requires collaborative ecosystems linking providers, vendors, regulators, and researchers to advance interoperability standards [27]. Standardized protocols such as DICOM, HL7, and FHIR, along with vendor-neutral archives (VNAs), are critical for seamless data exchange and reducing vendor lock-in [43]. Collaborative data curation aligned with clinical needs and clinician-led AI stewardship committees ensure practical applicability and smooth change management [10].

♦ **Developing Robust Governance and Regulatory Compliance Frameworks.** Comprehensive governance frameworks embedding ethics, privacy, and regulatory compliance are essential for mitigating risks such as bias, hallucinations, and data breaches [4, 52]. Continuous monitoring, validation, and predetermined change control plans support safe deployment under FDA and global regulatory guidance [20, 3]. Clear legal accountability and workforce training on safe AI use, HIPAA compliance, and MLLM limitations further strengthen adoption, trust, and operational reliability [63, 4].

♦ **Embedding Safety-by-Design Principles.** Safety must be a first-class design objective across all adjacent technologies, encompassing biosafety safeguards, fail-safe defaults, redundancy mechanisms, and real-time alerting for unsafe outputs. Proactive hazard analysis and resilience engineering approaches ensure that safety considerations are integrated from development through deployment [20].

♦ **Establishing Continuous Clinical Safety Monitoring.** Beyond technical validation, clinical MLLMs require continuous post-deployment safety surveillance, including incident reporting pipelines, safety audits, and clinician-in-the-loop feedback systems. Such monitoring reduces patient harm risks, ensures rapid error correction, and aligns with pharmacovigilance-style oversight models increasingly adopted in digital health [3, 4].

## 5 Concluding Remarks

The transformative potential of clinical Multimodal Large Language Models (MLLMs) resides not solely in their computational sophistication or cross-modal reasoning, but in the surrounding ecosystem that enables safe, reliable, and scalable deployment [12, 21]. In this position paper, we have systematically highlighted the strategic role of adjacent technologies, including specialized data curation pipelines, multimodal data lakes, secure API gateways, workflow integration middleware, continuous model monitoring platforms, and EHR/PACS connectors, as safety-critical enablers rather than peripheral tools. By integrating these components, organizations can construct resilient,

interoperable, secure, and auditable infrastructures that prioritize biosafety and patient safety, while mitigating risks such as data breaches, model hallucinations, and biased outputs [3, 27]. Furthermore, we have emphasized human-centered design, explainability, and clinician engagement as critical levers for trust, adoption, and workflow integration, thereby aligning technological capabilities with safe and responsible clinical practice [34]. Our analysis positions the *Clinical MLLM Adjacent* ecosystem as a distinct, rapidly evolving sector, underscoring the infrastructure premium in healthcare AI and the necessity for safety-driven strategic investment and collaborative development [39].

Looking forward, realizing the full diagnostic, therapeutic, and operational potential of MLLMs requires a systems-level approach that combines integrated infrastructure, governance, and workforce preparedness [4, 63]. Continuous model monitoring, robust regulatory alignment, and cross-sector standardization will ensure ethical, compliant, and scalable deployment, while fostering clinician confidence in AI-driven decision support. Importantly, the adoption of these adjacent technologies facilitates not only technical feasibility but also cultural transformation within healthcare organizations, enabling human-AI symbiosis and sustainable integration [34, 5]. By delineating this roadmap, we provide a structured framework for stakeholders to strategically invest, implement, and govern clinical MLLMs, thereby transforming these models from isolated innovations into operationally integrated, high-impact tools at the point of care. Collectively, this position reinforces the argument that the future of clinical AI hinges on the intelligent orchestration of MLLMs and their enabling ecosystem, setting the stage for equitable, precise, and scalable healthcare delivery.

## References


[1] Gwénolé Abgrall, Andre L. Holder, Zaineb Chelly Dagdia, Karine Zeitouni, and Xavier Monnet. Should AI models be explainable to clinicians? *Critical Care*, 28(1):301, September 2024.

[2] Evidently AI. Model monitoring framework. https://www.evidentlyai.com/ml-in-production/model-monitoring, 2025. Accessed July 1, 2025.

[3] Fiddler AI. Model monitoring framework. https://www.fiddler.ai/ml-model-monitoring/model-monitoring-framework, 2025. Accessed July 1, 2025.

[4] Simbo AI. Artificial intelligence in healthcare: Navigating regulatory challenges and ensuring compliance. https://www.simbo.ai/blog/artificial-intelligence-in-healthcare-navigating-regulatory-challenges-and-ensuring-compliance-262580/, 2025. Accessed July 1, 2025.

[5] Simbo AI. The future of healthcare jobs: Understanding the impact of automation and ai on employment in the medical field. https://www.simbo.ai/blog/the-future-of-healthcare-jobs-understanding-the-impact-of-automation-and-ai-on-employment-in-the-medical-field-32483/, 2025. Accessed July 14, 2025.

[6] Simbo AI. Investment trends in healthcare technology: Analyzing the growth of generative ai and its future implications. https://www.simbo.ai/blog/investment-trends-in-healthcare-technology-analyzing-the-growth-of-generative-ai-and-its-future-implications-2954736/, 2025. Accessed July 14, 2025.

[7] Rawan AlSaad, Alaa Abd-alrazaq, Sabri Boughorbel, Arfan Ahmed, Max-Antoine Renault, Rafat Damseh, and Javaid Sheikh. Multimodal large language models in health care: Applications, challenges, and future outlook. *Journal of Medical Internet Research*, 26:e59505, September 2024.

[8] CFlow Apps. Ai in clinical workflows: How healthcare providers can improve efficiency and patient care. https://www.cflowapps.com/ai-for-clinical-workflows/, 2025. Accessed July 14, 2025.

[9] Gopalakrishnan Arjunan. Implementing explainable ai in healthcare: Techniques for interpretable machine learning models in clinical decision-making. *International Journal of Scientific Research and Management (IJSRM)*, 9:597–603, 05 2021.

[10] Gaurav Belani. Deploying ai models in clinical workflows: Challenges and best practices. https://www.dataversity.net/deploying-ai-models-in-clinical-workflows-challenges-and-best-practices/, 2025. Accessed July 1, 2025.

[11] FDA Law Blog. Small change: Fda's final predetermined change control plan (pccp) guidance ditches ml and adds some details, but otherwise sticks closely to the draft. https://www.thefdalawblog.com/2025/02/small-change-fdas-final-predetermined-change-control-plan-pccp-guidance-ditches-ml-and-adds-some-details-but-otherwise-sticks-closely-to-the-draft/, 2025. Accessed July 14, 2025.

[12] Tenyks Blogger. Multimodal large language models (mllms): Transforming computer vision, 2024. Accessed: July 17, 2025.

[13] Google Cloud. Overview of the cloud healthcare api. https://cloud.google.com/healthcare-api/docs/introduction, 2025. Accessed July 14, 2025.

[14] Google Cloud. Using apigee api management for ai. https://cloud.google.com/blog/products/api-management/using-apigee-api-management-for-ai, 2025. Accessed July 14, 2025.

[15] Cognome. Explainerai™ analytics for explainable ai governance. https://cognome.com/explainerai-analytics-for-explainable-ai-governance, 2025. Accessed July 1, 2025.

[16] Contextflow. Ai interest group for imaging (aigi) task force. https://contextflow.com/2023/08/07/ai-interest-group-for-imaging-aigi-task-force/, 2023. Accessed July 14, 2025.

[17] NVIDIA Corporation. Medtronic and nvidia collaborate to build ai platform for medical devices, 2023. Accessed July 17, 2025.

[18] Dataloop.ai. The hl7 fhir model v1 — 107 different entities. https://dataloop.ai/library/model/sandeepkanao_hl7-fhir-model-v1, 2025. Accessed July 14, 2025.

[19] Elucidata. Streamline multi-modal data management in healthcare with ai. https://www.elucidata.io/blog/multi-modal-data-management-in-healthcare-strategies-for-integration-and-overcoming-data-silos, 2025. Accessed July 1, 2025.

[20] FDA. Artificial intelligence-enabled medical devices. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices, 2025. Accessed July 14, 2025.

[21] David A Feldstein, Isabel Barata, Thomas McGinn, Emily Heineman, Joshua Ross, Dana Kaplan, Francesca Bullaro, Sundas Khan, Nicholas Kuehnel, and Rachel P Berger. Disseminating child abuse clinical decision support among commercial electronic health records: Effects on clinical practice. *JAMIA Open*, 6(2):ooad022, April 2023.

[22] Xuelu Feng, Yunsheng Li, Dongdong Chen, Mei Gao, Mengchen Liu, Junsong Yuan, and Chunming Qiao. Benchmarking large and small mllms, 2025.

[23] Market Research Future. Ai clinical care market size, industry growth report 2034. https://www.marketresearchfuture.com/reports/ai-clinical-care-market-29087, 2025. Accessed July 14, 2025.

[24] G2. Best healthcare integration engines: User reviews from july 2025. https://www.g2.com/categories/healthcare-integration-engines, 2025. Accessed July 14, 2025.

[25] GlobeNewswire. Artificial intelligence (ai) in healthcare research report 2025 – global forecast to 2030. https://www.globenewswire.com/news-release/2025/06/04/3093732/28124/en/Artificial-Intelligence-AI-in-Healthcare-Research-Report-2025-Global-Forecast-to-2030-Key-Players-Expanding-their-Market-Share-Through-Product-Launches-and-Partnerships.html, 2025. Accessed July 14, 2025.

[26] Xianchao Guan, Zheng Zhang, Yifeng Wang, and Yongbing Zhang. *A systematic review on Multimodal Large Language Models (MLLMs) in computational pathology*. January 2025.

[27] Human Integrity HR. Ehr imaging integration challenges and practical solutions. https://www.humanintegrityhr.com/post/integrating-imaging-systems-with-ehrs-challenges-and-solutions, 2025. Accessed July 14, 2025.

[28] Alation Inc. Data curation. https://www.alation.com/glossary/data-curation/, 2025. Accessed July 1, 2025.

[29] Core Mobile Inc. General 1 — core mobile integrated workflow manager. https://www.coremobileinc.com/integrated-workflow-manager, 2025. Accessed July 14, 2025.

[30] Snowflake Inc. Healthcare & life sciences data cloud. https://www.snowflake.com/en/solutions/industries/healthcare-and-life-sciences/, 2025. Accessed July 1, 2025.

[31] SNS Insider. Healthcare middleware market size, share & growth report 2032. https://www.snsinsider.com/reports/healthcare-middleware-market-6078, 2025. Accessed July 14, 2025.

[32] Business Research Insights. Pacs systems market size, share, trends, growth report, 2033. https://www.businessresearchinsights.com/market-reports/pacs-systems-market-122721, 2025. Accessed July 14, 2025.

[33] Global Growth Insights. Departmental pacs market insights report 2033. https://www.globalgrowthinsights.com/market-reports/departmental-picture-archiving-and-communication-system-pacs-market-103062, 2025. Accessed July 14, 2025.

[34] Eric Karofsky. Beyond the molecule: How human-centered design unlocks ai's promise in pharma. https://www.appliedclinicaltrialsonline.com/view/human-centered-ai-pharma, 2025. Accessed July 1, 2025.

[35] Keragon. The economic impact of ai in healthcare: Key considerations for 2025 and beyond. https://www.keragon.com/blog/economic-impact-of-ai-in-healthcare, 2025. Accessed July 14, 2025.

[36] Dany Kitishian. Astrazeneca's ai strategy: Analysis of ai dominance in pharmaceutical and biotech. https://www.klover.ai/astrazeneca-ai-strategy-analysis-of-ai-dominance-in-pharmaceutical-biotech/, July 2025. Accessed July 17, 2025.

[37] KrakenD. Api gateway for llms. https://www.krakend.io/docs/ai-gateway/, 2025. Accessed July 14, 2025.

[38] MLM Medical Labs. Real time clinical trial data – mlm online. https://www.mlm-labs.com/services/mlm-online/, 2025. Accessed July 14, 2025.

[39] Heather Landi. Healthcare ai rakes in nearly $4b in vc funding, buoying the digital health market in 2025. https://www.fiercehealthcare.com/health-tech/healthcare-ai-rakes-nearly-4b-vc-funding-buoying-digital-health-market-2025, July 2025. Accessed July 1, 2025.

[40] HongYi Li, Jun-Fen Fu, and Andre Python. Implementing large language models in health care: Clinician-focused review with interactive guideline. *J Med Internet Res*, 27:e71916, Jul 2025.

[41] Zeyu Liu, Zhitian Hou, Yining Di, Kejing Yang, Zhijie Sang, Congkai Xie, Jingwen Yang, Siyuan Liu, Jialu Wang, Chunming Li, Ming Li, and Hongxia Yang. Infi-med: Low-resource medical mllms with robust reasoning evaluation, 2025.

[42] AZ Big Media. Why digital health is the next big investment trend in arizona's medical industry. https://azbigmedia.com/business/health-care/why-digital-health-is-the-next-big-investment-trend-in-arizonas-medical-industry/, July 2025. Accessed July 14, 2025.

[43] Medicai. Pacs integration for faster diagnoses & better patient care. https://blog.medicai.io/en/pacs-integration-and-workflow/, 2025. Accessed July 14, 2025.

[44] Medicai. Structured radiology reporting – ai application for enhanced operation. https://blog.medicai.io/en/structured-radiology-reporting-ai-application/, 2025. Accessed July 14, 2025.

[45] Mindbowser. Multimodal ai in healthcare: Powering smart care systems. https://www.mindbowser.com/multimodal-ai-in-healthcare/, 2025. Accessed July 1, 2025.

[46] Aya El Mir, Lukelo Thadei Luoga, Boyuan Chen, Muhammad Abdullah Hanif, and Muhammad Shafique. Democratizing mllms in healthcare: Tinyllava-med for efficient healthcare diagnostics in resource-constrained settings, 2024.

[47] Jose Gabriel Islas Montero and Dmitry Kazhdan. Multimodal large language models (mllms) transforming computer vision. *The Tenyks Blogger*, 2024. Accessed: July 14, 2025.

[48] PR Newswire. Artificial intelligence in diagnostics market size worth us$ 5.44 billion by 2030 - exclusive report by the research insights. https://www.prnewswire.com/news-releases/artificial-intelligence-in-diagnostics-market-size-worth-us-5-44-billion-by-2030—exclusive-report-by-the-research-insights-302461670.html, 2025. Accessed July 14, 2025.

[49] Northeastern Online. Ai's role in healthcare for non-technical professionals. https://online.northeastern.edu/resources/ai-in-healthcare/, 2025. Accessed July 14, 2025.

[50] Orases. Healthcare software: Apis and middleware platforms. https://orases.com/blog/apis-middleware-platforms-for-healthcare/, 2025. Accessed July 14, 2025.

[51] Saurabh Pahune, Zahid Akhtar, Venkatesh Mandapati, and Kamran Siddique. The importance of ai data governance in large language models, 04 2025.

[52] Saurabh Pahune, Zahid Akhtar, Venkatesh Mandapati, and Kamran Siddique. The importance of ai data governance in large language models. *Big Data and Cognitive Computing*, 9(6), 2025.

[53] Saigurudatta Pamulaparthyvenkata. Applications of multimodal large language models in personalized healthcare. *Towards AI*, 2024. Accessed: July 14, 2025.

[54] Purview. Ehr pacs integration | connect your ehr with medical images. https://www.purview.net/connect-your-ehr-with-medical-images, 2025. Accessed July 14, 2025.

[55] Jianing Qiu, Wu Yuan, and Kyle Lam. The application of multimodal large language models in medicine. *The Lancet Regional Health – Western Pacific*, 45, April 2024. Publisher: Elsevier.

[56] Radsource. Dicom vs hl7 – everything you need to know. https://radsource.us/dicom-vs-hl7/, 2025. Accessed July 14, 2025.

[57] Dimension Market Research. Healthcare api market set to hit usd 343.8 million by 2033. https://dimensionmarketresearch.com/report/healthcare-api-market/, 2025. Accessed July 14, 2025.

[58] Precedence Research. Ai-based clinical trials solution provider market size to hit usd 17.40 bn by 2034. https://www.precedenceresearch.com/ai-based-clinical-trials-solution-provider-market, 2025. Accessed July 14, 2025.

[59] Straits Research. Healthcare api market size, share, industry trends & forecast to 2033. https://straitsresearch.com/report/healthcare-api-market, 2025. Accessed July 14, 2025.

[60] Amazon Web Services. Healthlake partners. https://aws.amazon.com/healthlake/partners/, 2025. Accessed July 1, 2025.

[61] Amazon Web Services. Multi-modal, multi-omics data integration & analysis on aws. https://aws.amazon.com/health/solutions/mmmo/, 2025. Accessed July 1, 2025.

[62] Shaip. Unlocking healthcare ai potential with multimodal medical datasets. https://www.shaip.com/blog/multimodal-medical-datasets-for-ai-research/, March 2025. Accessed July 1, 2025.

[63] TechMagic. Hipaa compliance ai: Guide to using llms safely in healthcare. https://www.techmagic.co/blog/hipaa-compliant-llms, 2025. Accessed July 1, 2025.

[64] Manoj Vallikkat. Ai-powered healthcare in asia pacific: What's next for 2025 and beyond? https://blogs.idc.com/2025/07/11/ai-powered-healthcare-in-asia-pacific-whats-next-for-2025-and-beyond/, 2025. Accessed July 1, 2025.

[65] Yiqi Wang, Wentao Chen, Xiaotian Han, Xudong Lin, Haiteng Zhao, Yongfei Liu, Bohan Zhai, Jianbo Yuan, Quanzeng You, and Hongxia Yang. Exploring the reasoning abilities of multimodal large language models (mllms): A comprehensive survey on emerging trends in multimodal reasoning, 2024.

[66] Honglong Yang, Shanshan Song, Yi Qin, Lehan Wang, Haonan Wang, Xinpeng Ding, Qixiang Zhang, Bodong Du, and Xiaomeng Li. Multi-modal explainable medical ai assistant for trustworthy human-ai collaboration, 2025.

[67] Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, and Enhong Chen. A survey on multimodal large language models. *National Science Review*, 11(12), November 2024.

[68] Tomoko Yokoi and Michael R. Wade. Strategic use of ai in healthcare and pharma drives market advantage. here's how. https://www.imd.org/ibyimd/artificial-intelligence/strategic-use-of-ai-in-healthcare-and-pharma-drives-market-advantage-heres-how/, July 2025. Accessed July 1, 2025.

[69] Kaiwen Zuo and Yirui Jiang. Medhallbench: A new benchmark for assessing hallucination in medical large language models, 2025.

# A Supplementary Material

Table 1: Market Size and Growth Projections for Key Clinical AI Infrastructure Segments (2024–2034)

| **Market Segment** | **2024/2025 Market Size ($B)** | **Projected 2030-2032-2034 Market Size ($B)** | **CAGR (%)** |
|---|---|---|---|
| AI in Healthcare (Overall) | $21.66 (2025) | $110.61 (2030) | 38.6% |
| AI in Clinical Trials | $2.88 (2025) | $17.40 (2034) | 22.13% |
| AI in Diagnostics | $1.97 (2025) | $5.44 (2030) | 22.46% |
| AI Clinical Care | $11.35 (2025) | $95.15 (2034) | 26.65% |
| Healthcare API Market | $1.38 (2025) | $1.92 (2033) | 4.2% |
| Healthcare Middleware Market | $4.52 (2024) | $8.68 (2032) | 7.52% |
| PACS Systems Market | $5.41 (2024) | $7.601 (2033) | 3.8% |
| Healthcare Data Monetization Market | $0.62 (2025) | $1.19 (2030) | 14.10% |

Table 2: Key Regulatory and Ethical Considerations for Clinical MLLM Deployment

| **Consideration Area** | **Key Regulations/ Guidelines** | **Impact on MLLM Deployment** | **Adjacent Technology Role** |
|---|---|---|---|
| Data Privacy and Security | HIPAA, GDPR, WHO Ethics Guidance | Risk of data breaches, unauthorized access, misuse of PHI. | Secure API Gateways, Data Curation (de-identification), Multimodal Data Lakes (secure storage) |
| Algorithmic Bias and Fairness | WHO Ethics Guidance, FDA SaMD principles, EU AI Act | Exacerbating health disparities, unfair treatment, loss of public trust. | Data Curation (diverse datasets), Model Monitoring (bias detection), Explainable AI (fairness analysis) |
| Transparency and Explainability | EU GDPR (*right to explanation*), FDA SaMD principles, EU AI Act | Lack of clinician trust, difficulty in validating decisions, unclear accountability. | Model Monitoring (XAI platforms, saliency maps, SHAP values), Workflow Integration Middleware (context-aware alerts) |
| Safety and Effectiveness Monitoring | FDA SaMD/PCCP, ISO 13485 | Potential for patient harm (e.g., hallucinations), model degradation over time, need for continuous validation. | Model Monitoring (drift detection, hallucination detection, performance monitoring), Data Curation (ground truth updates) |
| Medical Liability | Evolving legal frameworks | Unclear accountability for AI-driven errors, increased legal risk for providers/developers. | Secure API Gateways (audit logging), Data Governance (provenance, clear policies), Model Monitoring (performance logs) |
| Continuous Learning/Adaptive AI | FDA Predetermined Change Control Plans (PCCP) | Challenges with post-market modifications, ensuring ongoing safety and effectiveness, regulatory approval for updates. | Model Monitoring (re-training practices, performance evaluation protocols), Data Curation (high-quality SFT datasets) |

# B More Details on Pillars of the Clinical MLLM Ecosystem

The effective integration of MLLMs into clinical practice is contingent upon a coordinated ecosystem of adjacent technologies (Table 3), encompassing data lakes, workflow middleware, secure API gateways, monitoring frameworks, and interoperable connectors to EHR and PACS systems. Together, these pillars constitute the infrastructural backbone that enables MLLMs to generate clinically accurate, reliable, and ethically compliant outputs, while supporting real-time deployment, continuous model monitoring, and adherence to regulatory and privacy standards. By embedding these technologies into end-to-end healthcare workflows, organizations can bridge the gap between model capabilities and actionable clinical decision-making, ensuring that the transformative potential of MLLMs translates into tangible patient benefit.

## B.1 Specialized Data Curation Tools

Data curation stands as a foundational pillar in the healthcare AI ecosystem, particularly in the development and deployment of Multimodal Large Language Models (MLLMs). Given the sheer volume and heterogeneity of data originating from patient records, clinical trials, imaging archives, and biomedical literature, ensuring data quality, consistency, and integrity is essential [28]. The process of data curation not only eliminates errors, duplicates, and inconsistencies but also upholds data provenance, crucial for ensuring transparency, reproducibility, and accountability throughout the MLLM lifecycle [28].

At the core of effective curation are sophisticated capabilities for de-identification and annotation. These tools are essential for transforming raw data into usable, privacy-compliant training resources. De-identification systems must rigorously anonymize Protected Health Information (PHI) across both textual and visual modalities,

Table 3: Core Functions of Clinical MLLM Adjacent Technologies

| Technology Category | Key Functions | Role in MLLM Success | Example Vendors/Standards |
|---|---|---|---|
| Specialized Data Curation Tools | Data quality, integrity, provenance, de-identification, annotation, bias mitigation. | Fuels accurate, unbiased MLLM training; ensures ethical data use. | Shaip, Elucidata Polly |
| Multimodal Data Lakes | Unification of diverse structured, unstructured, and streaming data; scalable storage, querying. | Enables holistic patient insights for MLLMs; supports complex multimodal reasoning. | AWS HealthLake, Snowflake |
| Robust Model Monitoring Platforms | Detects data drift, model degradation, anomalies; identifies and mitigates hallucination and bias; provides Explainable AI (XAI). | Ensures continuous performance, safety, and trustworthiness of MLLMs in production. | Fiddler AI, Evidently AI, Cognome ExplainerAI™ |
| Secure API Gateways | Authentication, authorization, traffic management, cost control, prompt validation, policy enforcement. | Provides secure, compliant, and scalable access to MLLMs and sensitive data. | Google Cloud Healthcare API, Apigee, KrakenD |
| Workflow Integration Middleware | Bridges legacy systems with modern AI; automates tasks, optimizes operational efficiency, reduces cognitive load; orchestrates processes. | Seamlessly embeds MLLMs into existing clinical and administrative workflows; enhances clinician adoption. | Core Mobile PCSIP, NextGen Mirth, Cflow, Lionbridge Aurora AI |
| Specialized EHR/PACS Connectors | Facilitates real-time, standardized access to electronic health records and medical images; enables AI-driven structured reporting. | Provides the comprehensive clinical data backbone for MLLM training and inference; transforms clinical documentation. | DICOM, HL7, FHIR, Medicai, Purview |

ensuring compliance with healthcare regulations such as HIPAA while retaining the clinical utility of the data [62, 63]. In parallel, high-quality annotation, often performed by medical experts, enables the construction of structured, labeled datasets that are critical for supervised learning and fine-tuning. Recent advancements also incorporate instruction augmentation and chain-of-thought annotations that enrich multimodal datasets, boosting the domain-specific reasoning and cross-modal integration capacity of MLLMs [41]. The influence of curated datasets on MLLM performance cannot be overstated. High-fidelity, well-annotated data is indispensable for model training, fine-tuning, and post-deployment refinement. It plays a direct role in reducing hallucinations, improving factual accuracy, and enhancing clinical plausibility in AI-generated outputs [52, 69]. Critically, data curation also functions as a primary mechanism for addressing algorithmic bias. Biased training data can perpetuate systemic health disparities, especially when models are deployed across diverse populations. Through careful dataset construction, including the selection of balanced and representative samples, data curation mitigates this risk and aligns with broader ethical imperatives in AI development [53, 52].

Thus, data curation is far more than a backend process, it is a strategic endeavor that transforms fragmented and often siloed medical data into an ethically compliant, clinically relevant, and AI-ready asset. This transformation is pivotal in enabling MLLMs to operate safely and effectively in healthcare environments, where the cost of error can be profound. The ongoing nature of curation ensures that datasets evolve alongside clinical standards and regulatory requirements, reinforcing the trustworthiness and long-term utility of MLLMs [51]. Several industry leaders exemplify the state of the art in this space. Shaip, for instance, offers extensive pre-processed datasets, expert annotation and labeling services, and robust de-identification solutions for multimodal medical data [62]. Elucidata's Polly platform provides an end-to-end solution, including centralized data ingestion, metadata harmonization, and scalable annotation engines, all within a secure and compliant framework [19]. These platforms illustrate how specialized data curation tools are becoming critical enablers of MLLM accuracy, trust, and clinical applicability.

### B.2 Multimodal Data Lakes

Multimodal data lakes form the critical architectural backbone for MLLMs in healthcare, designed to unify and manage the vast and heterogeneous datasets that characterize clinical practice. These platforms seamlessly integrate structured data such as Electronic Health Records (EHRs) and laboratory results with unstructured and streaming data, including medical images, clinical notes, audio recordings, video feeds, and omics datasets, all within a secure and compliant environment [7]. This unified approach facilitates a holistic, longitudinal view of an individual's health, moving beyond fragmented information silos to generate richer insights that span from research settings to bedside care [61]. By aggregating diverse sources, ranging from diagnostic imaging and wearable sensor data to genetic profiles and patient-reported outcomes, multimodal data lakes empower MLLMs to achieve deeper, context-aware understanding [45]. This capability mirrors the integrative reasoning clinicians employ when diagnosing and treating patients, enabling models to correlate findings across modalities: for example, linking CT scan results with pathology reports, combining audio from telehealth sessions with clinical documentation, or merging continuous sensor streams from smartwatches or glucose monitors with EHR data [7]. Such integration is essential for advancing personalized medicine and refining predictive analytics, ultimately supporting more precise treatment planning and earlier disease detection [53].

**Scalability and Security: Foundations for Clinical Use.** The scalability and secure storage capabilities of multimodal data lakes are paramount to their success. Leading cloud-based platforms such as AWS and Snowflake provide robust infrastructure to store multi-terabyte datasets in centralized repositories, significantly reducing data silos and enabling near real-time access to cross-modal information [19]. These platforms offer elastic computational resources that can support parallel processing of extensive omics and clinical datasets while ensuring compliance with healthcare regulations through certifications such as HITRUST and adherence to HIPAA standards [19]. More than mere storage solutions, multimodal data lakes are the essential foundation that allows MLLMs to realize their full *multimodal* potential. Traditional data warehouses or disconnected siloed systems are ill-equipped to meet the dynamic and complex data demands of these models. Data lakes are purpose-built to harmonize and interconnect disparate data points, creating a unified, intelligent view of patient information that closely parallels clinical decision-making processes [45]. This capability extends beyond data availability to encompass data interoperability and accessibility, critical enablers for sophisticated AI analysis. The concept of *any-to-any MLLMs*, capable of ingesting and generating outputs across all modalities, fundamentally depends on data lakes' ability to ingest, harmonize, and present diverse data types in an AI-consumable format. Consequently, data readiness emerges as a primary bottleneck in MLLM deployment, often outweighing challenges related to model architecture or training.

Key providers in this space include AWS for Health, which offers specialized services like AWS HealthOmics for genomic data and AWS HealthImaging for medical imaging, complemented by an ecosystem of partner solutions designed to unify and analyze multimodal healthcare data [61]. Additionally, AWS HealthLake partners facilitate the transformation of legacy healthcare data into standardized formats such as Fast Healthcare Interoperability Resources (FHIR), while providing tools for efficient health record navigation and visualization [60]. Snowflake's AI Data Cloud similarly delivers a unified platform capable of handling unstructured, semi-structured, and structured data, enabling real-time patient insights and secure collaboration across healthcare networks [30]. Together, these multimodal data lakes lay the indispensable groundwork for clinical MLLMs, supporting their integration into healthcare workflows and unlocking new horizons in precision medicine.

### B.3 Robust Model Monitoring Platforms

Robust model monitoring platforms are essential for the safe and effective clinical deployment of MLLMs. Continuous tracking of model performance and behavior in production ensures sustained accuracy, reliability, and safety over time [3]. Key functionalities include detecting data drift, shifts in input distributions, model degradation, and anomalies that may impair performance [3]. These platforms provide actionable insights that allow developers to diagnose issues promptly and maintain optimal model functioning [3].

A critical focus of MLLM monitoring in healthcare is managing hallucinations and bias. MLLMs have a known tendency for *hallucinations*, producing medically implausible or inaccurate outputs that risk patient safety by leading to misdiagnoses or inappropriate treatments [69]. Monitoring platforms assess these risks using expert-validated case scenarios and systematic annotation methods that categorize hallucinations by anatomical and pathological types [69]. Additionally, they help detect and mitigate bias arising from training data imbalances or fine-tuning, ensuring equitable model behavior across diverse patient populations [52].

**Explainability and Safety as Core Monitoring Pillars.** Explainable AI (XAI) plays an integral role in effective monitoring by unraveling the *black box* nature of complex MLLMs, making their decisions interpretable and trustworthy to clinicians [9]. Techniques such as feature attribution methods (e.g., SHAP values, saliency maps) and example-based explanations facilitate clinician understanding and confidence in AI outputs [9]. This transparency is crucial for patient safety, regulatory compliance, such as the EU GDPR *right to explanation*, and fostering robust human-AI collaboration [9]. While some advocate prioritizing accuracy over explainability in urgent clinical scenarios, the consensus for MLLMs emphasizes transparency to build and sustain trust [1]. Beyond traditional performance metrics like accuracy and precision, clinical MLLM monitoring prioritizes patient safety outcomes. Monitoring must detect specific instances of *anatomical* and *pathological hallucinations*, quantifying their potential harm to patients [69]. This elevates monitoring from a technical exercise to a vital component of clinical governance, risk management, and ethical AI deployment [9]. The increasing regulatory focus on XAI and bias mitigation further underscores this transformation [9].

Leading solutions in this space include Fiddler AI's Observability platform, offering performance monitoring, drift detection, quality assurance, custom alerts, and specialized NLP and computer vision monitoring [3]. Evidently AI provides an open-source library supporting issue detection, root cause, and behavioral analysis for ML models [2]. Cognome's ExplainerAI™ specializes in healthcare AI transparency, integrating with EHRs to assist in bias detection and regulatory compliance [15]. Additionally, platforms like MLM-Labs and Biologit address healthcare-specific needs by enabling real-time monitoring of clinical trial data and medical literature for safety surveillance [38]. Together, these model monitoring platforms form an indispensable safeguard that ensures MLLMs maintain clinical performance, uphold patient safety, and comply with evolving ethical and regulatory standards.

### B.4 Secure API Gateways

Secure API gateways are foundational to the integration and deployment of MLLMs in healthcare, acting as specialized control layers for AI workloads. They provide secure, scalable, and manageable access to MLLMs and related AI services, an imperative in an industry handling highly sensitive patient data [14]. These gateways enforce zero-trust security principles, manage authentication and authorization, and serve as critical checkpoints for all MLLM interactions, safeguarding sensitive data throughout [63]. These platforms deliver a comprehensive suite of features to enhance security, scalability, and governance. They facilitate seamless AI agent integration by offering robust authentication, traffic management, detailed analytics, and policy enforcement [14]. Additionally, API gateways can intelligently route requests to the most appropriate MLLM based on task requirements, optimize response times through semantic caching, and enforce granular usage limits to control inference costs [14].

Compliance with stringent healthcare data security standards, such as HIPAA in the U.S., is non-negotiable. API gateways play a pivotal role in maintaining compliance by implementing strong encryption for data at rest and in transit, enforcing strict role-based access controls (RBAC), and maintaining exhaustive logs of all interactions [63]. They also support de-identification of Protected Health Information (PHI), creating a secure, auditable layer that exposes sensitive ePHI to patient and provider applications while mitigating the risk of breaches and unauthorized access [13].

**API Gateways as Trust Boundaries.** In healthcare's highly regulated environment, API gateways function not just as technical interfaces but as critical *trust boundaries* that enforce security, compliance, and ethical AI use at every interaction point. This elevates MLLM access from a technical challenge to a governance imperative. While APIs generally facilitate data exchange [13], an *AI Gateway* specifically adds a tailored control layer for AI workloads, actively managing how MLLMs access and process sensitive PHI [37]. It enforces zero-trust policies, prompt validation, and compliance checks, operationalizing the trust boundary concept [37]. By enabling encryption, RBAC, and detailed audit logging, these gateways empower healthcare organizations to transition from passive regulatory compliance to proactive risk mitigation at the integration layer [63].

Leading providers in this domain include Google Cloud Healthcare API, which supports industry-standard protocols like DICOM, FHIR, and HL7v2 for ingesting, storing, and analyzing healthcare data. It is built on a robust security framework featuring Identity and Access Management (IAM) and comprehensive auditability via Cloud Logging, and is covered under Google Cloud's HIPAA Business Associate Addendum (BAA) [13]. Additional examples of API gateways offering advanced AI safety features, latency optimization, cost control, and governance for LLM interactions include Apigee (part of Google Cloud's API Management) and KrakenD [14, 37]. Together, secure API gateways constitute a critical infrastructure layer that enables the safe, compliant, and efficient integration of MLLMs into healthcare environments, balancing performance demands with stringent regulatory and ethical requirements.

### B.5 Workflow Integration Middleware

Workflow integration middleware is vital for the practical adoption of MLLMs in healthcare, serving as a bridge that enables seamless and secure data exchange between diverse healthcare software systems. This includes connecting MLLMs with existing EHRs, PACS, and medical devices [50]. Such middleware supports the incremental modernization of IT ecosystems, allowing providers to link legacy systems with cloud-based services without costly infrastructure overhauls [50]. A major advantage of middleware is its capacity to automate tasks, optimize efficiency, and reduce cognitive burden on healthcare professionals. AI-driven workflow automation is a priority for providers aiming to enhance operational efficiency and care quality [64]. Middleware enables MLLMs to automate repetitive, data-heavy tasks like appointment scheduling, billing, and documentation, thereby alleviating administrative workload, improving resource use, and reducing clinician burnout [64, 49]. Many hospitals still rely on legacy systems incompatible with modern AI or cloud environments [43]. Middleware effectively bridges this gap by converting file types, managing data transfers, and handling complex security protocols, minimizing disruptive full-system replacements [27]. This ensures smooth data flow and interoperability across clinical and administrative domains [50].

**From Integration to Orchestration.** Workflow integration middleware transforms MLLMs from isolated *AI tools* into components of cohesive, *AI-orchestrated workflows*, acting as the conductor that ensures seamless data exchange, task automation, and proactive care delivery [8]. Beyond simple integration, middleware coordinates tasks, monitors performance, and routes assignments based on real-time data and rules, enabling MLLM outputs to trigger downstream actions [8]. This intelligent orchestration addresses operational inefficiencies and clinician resistance, reduces documentation time, and directly mitigates physician burnout, key factors for sustainable AI adoption [42, 34]. Leading providers include Core Mobile's Patient Care Systems Integration Platform (PCSIP), which unifies diverse patient records and enables department-specific LLM customization [29]. Orases highlights API- and middleware-driven seamless data transfer within healthcare [50]. NextGen's Mirth Integration Engine standardizes clinical data flow and offers advanced alerting for system monitoring [24].

Additionally, platforms like Cflow and Lionbridge Aurora AI provide sophisticated workflow orchestration to integrate AI tools for clinical and administrative automation [8].

### B.6 Specialized EHR/PACS Connectors

Specialized EHR/PACS connectors serve as the vital clinical data backbone for MLLMs, enabling real-time, standardized access to electronic health records and medical imaging. These connectors provide MLLMs with immediate access to comprehensive patient data, including high-resolution medical images in DICOM format and extensive clinical records in HL7 and FHIR standards [27]. This seamless integration prevents workflow disruptions by allowing physicians to access imaging studies directly within patient records without navigating separate systems [43].

**Interoperability.** Interoperability remains a major challenge in healthcare IT, stemming from diverse vendor systems and heterogeneous data formats [27]. EHR/PACS connectors address these issues by leveraging established protocols: DICOM standardizes storage and transfer of medical images [27], HL7 governs exchange of clinical and administrative data [27], and FHIR, a modern, web-based standard with modular resources and RESTful APIs, facilitates flexible, semantic data exchange critical for multimodal integration [56]. Adoption of these standards creates a common language for effective communication between imaging and EHR systems [27]. Initiatives such as IHE integration profiles (e.g., AIW-I, AIR) further promote standardized AI interaction with DICOM data, reflecting ongoing efforts toward interoperability [16].

**Enabling Intelligent AI-Driven Clinical Workflows.** By granting MLLMs integrated access to imaging and clinical data, these connectors empower advanced applications like structured radiology reporting. This automation analyzes imaging (e.g., MRI, X-ray) and populates structured templates, reducing manual entry and error risk [44]. It enhances diagnostic accuracy by highlighting critical findings and supports clinicians in informed decision-making, while preserving clinician oversight [44]. Additionally, MLLMs can utilize this data to generate patient-friendly communications, improving engagement and understanding [44]. The evolution of EHR/PACS connectors marks a shift from passive repositories to active, intelligent participants within MLLM workflows. They enable continuous data exchange, where MLLMs receive clinical data and feed back AI-generated insights, such as reports, findings, and administrative elements like billing codes, directly into EHR/PACS [44]. This bidirectional flow, powered by robust connectors adhering to standards like FHIR, forms the foundation for truly *AI-native* clinical documentation and decision support. It transitions healthcare from manual data entry and fragmented records toward automated, integrated, and intelligent patient data ecosystems.

Notable solutions include Medicai's PACS integration connecting radiology and imaging departments with EHR/RIS systems via DICOM and HL7, enabling instant access and workflow optimization [43]. Purview offers EHR-PACS integration for seamless linkage between EHRs and medical images [54]. The Dataloop HL7 FHIR Model V1 demonstrates how AI models can leverage these integrated data formats to recognize biomedical entities in text, illustrating practical MLLM data utilization [18].

## C More Details on Solutions and Recommendations

To navigate the complexities and fully realize the potential of clinical MLLMs, a multi-faceted strategic approach is required, focusing on integrated infrastructure, human-centered design, collaboration, and robust governance.

### C.1 Strategic Investment in Integrated Infrastructure

Healthcare organizations and technology providers must prioritize strategic investments in the entire *MLLM adjacent* ecosystem, recognizing that isolated MLLMs offer limited value without robust supporting infrastructure. The market trends clearly indicate a significant *infrastructure premium* in digital health funding, with AI-enabled solutions and data infrastructure attracting the lion's share of investment [39]. This investment should comprehensively span specialized data curation tools for building high-quality, unbiased, and de-identified datasets essential for MLLM training and fine-tuning [52]. It must also include multimodal data lakes for unified, scalable access to diverse clinical data 28, and secure API gateways to ensure compliant and efficient data flow to and from MLLMs [14]. Developing comprehensive AI adoption roadmaps that explicitly budget for these adjacent technologies is crucial. Furthermore, fostering public-private partnerships can facilitate the co-development of shared infrastructure components, accelerating progress and reducing individual organizational burden. Investing heavily in cloud-based solutions is also recommended for their inherent scalability, cost-efficiency, and remote access capabilities, which are vital for managing the vast data volumes and computational demands of MLLMs [25].

### C.2 Prioritizing Human-Centered AI Design and Explainability

All adjacent technologies, particularly those interacting directly with clinicians and patients, must be designed with a human-centered approach, emphasizing explainability, usability, and trust. Clinician resistance and workflow disruption represent major impediments to AI adoption, often outweighing technical capabilities [34]. The success of AI in healthcare hinges on a positive human experience, necessitating transparent and intuitive interfaces that seamlessly integrate into existing clinical workflows [34]. Explainable AI (XAI) is critical for clinicians to understand and trust MLLM recommendations, which is fundamental for patient safety and adherence to regulatory mandates [9]. To achieve this, human-centered design principles should be implemented from the outset of development, ensuring that AI solutions augment, rather than complicate, clinical practice. Integrating XAI features into model monitoring platforms and directly into MLLM outputs will provide the necessary transparency. Furthermore, conducting silent trials and actively engaging clinicians early in the design and implementation phases is vital for mapping existing workflows, identifying pain points, and securing crucial buy-in from end-users [10].

### C.3 Cross-Stakeholder Collaboration and Standardization

Driving widespread adoption of interoperability standards and fostering collaborative ecosystems involving healthcare providers, technology vendors, regulators, and research institutions is essential. The lack of standardized communication protocols between disparate systems, such as EHRs and PACS, remains a significant hurdle to seamless MLLM integration [27]. Adopting standardized protocols like DICOM (for imaging), HL7 (for clinical data exchange), and particularly FHIR (for modern, web-based interoperability) is crucial for seamless data exchange and MLLM functionality [43].Encouraging the use of vendor-neutral archives (VNAs) can help overcome vendor lock-in and promote broader data accessibility [43].Collaboration ensures that data curation efforts align with real-world organizational needs and that AI solutions are developed with practical clinical applicability.5 Establishing AI stewardship committees with rotating clinician leadership can effectively guide implementation, manage change, and ensure that technological advancements are aligned with clinical realities and needs [10].

### C.4 Developing Robust Governance and Regulatory Compliance Frameworks

Establishing comprehensive AI data governance frameworks that embed ethical principles, privacy safeguards, and regulatory compliance throughout the entire MLLM lifecycle is non-negotiable. The sensitive nature of healthcare data and the adaptive, continuously learning nature of MLLMs necessitate strong governance to mitigate risks such as algorithmic bias, hallucinations, and data breaches [4]. Regulatory bodies, including the FDA, are increasingly focusing on post-market monitoring and predetermined change control plans for AI-enabled medical devices, underscoring the need for continuous oversight [20]. Implementing robust data governance policies that cover data quality, privacy (e.g., encryption, role-based access controls, de-identification), and ethical AI standards is paramount [52]. Organizations must develop internal policies for continuous monitoring and validation of MLLM performance in production environments to detect and address issues proactively [3]. Ensuring that legal accountability is clearly defined for AI-driven decisions is also critical for building trust and managing risk [4]. Finally, investing in comprehensive training programs for all staff on safe AI use, HIPAA compliance, and the limitations and strengths of MLLMs will empower the workforce and foster responsible adoption [63].

### C.5 Embedding Safety-by-Design Principles.

Safety should not be treated as an afterthought but embedded throughout the lifecycle of adjacent technologies. This includes incorporating biosafety safeguards during data ingestion, implementing fail-safe defaults in workflow integration middleware, and designing redundancy mechanisms to prevent single points of failure. Real-time alerting for anomalous or unsafe outputs is critical for clinical environments where delays can cause harm. Hazard analysis and resilience engineering approaches, such as Failure Mode and Effects Analysis (FMEA), can be adapted to MLLM deployment pipelines to anticipate and mitigate risks proactively. Embedding these principles ensures that infrastructures remain robust under uncertainty and maintain patient safety as a core operational priority [20].

### C.6 Establishing Continuous Clinical Safety Monitoring.

Clinical MLLMs require ongoing post-deployment safety surveillance that extends beyond technical validation. This involves establishing structured incident reporting pipelines for clinicians, conducting routine safety audits, and integrating clinician-in-the-loop feedback mechanisms within monitoring platforms. Borrowing from pharmacovigilance, continuous surveillance should track not only model drift but also real-world safety outcomes such as misdiagnoses, inappropriate recommendations, or workflow disruptions. Automated safety dashboards, coupled with human oversight, provide rapid detection and remediation of risks. Such an approach

ensures that clinical MLLMs evolve responsibly over time while minimizing patient harm and aligning with emerging global regulatory expectations [3, 4].