Integrative Survey of Multimodal Analytics, IoT-Enabled Biometric Monitoring, and Artificial Intelligence for Secure, Scalable, and Explainable Healthcare Systems

Abstract

This survey comprehensively examines the current landscape and future directions of automated, data-driven, and AI-enabled analytics in healthcare, addressing the surging complexity and heterogeneity of clinical, sensor, and IoT-derived data. Motivated by the imperatives of real-time monitoring, precision diagnostics, and personalized therapeutics, the review synthesizes methodological advances across multimodal data fusion, biomedical signal processing, deep learning, and explainable artificial intelligence (XAI). The scope encompasses foundational infrastructures, taxonomy of data modalities, benchmarking datasets, and emergent computational frameworks enabling scalable, privacy-preserving, and interoperable analytics. Key contributions include a critical evaluation of signal processing methods for real-time health monitoring, state-of-the-art multimodal and cross-modal fusion strategies, and adaptive analytics tailored for imperfect, high-velocity IoT streams. The survey elucidates the interplay between synthetic data generation and regulatory, ethical, and technical challenges, emphasizing the centrality of transparency, auditability, and robust privacy/security foundations—particularly within federated and resource-constrained environments. By integrating case studies, comparative paradigms, and application-driven insights, the review highlights best practices in operational analytics, clinical decision support, and the transition to proactive and participatory models of care. Concluding, it identifies persistent challenges—data heterogeneity, annotation scarcity, standardization, equity, interoperability, explainability, and responsible AI governance—charting a coherent roadmap for the transformative, scalable, and ethical deployment of AI-enabled analytics across the biomedical and health sector.

1 Introduction

Artificial Intelligence (AI) has achieved unprecedented milestones in recent years, permeating diverse domains such as natural language processing, computer vision, and decision support systems. The rapid evolution of algorithms, model architectures, and data resources continually redefines the state of the art, compelling both newcomers and practitioners to stay abreast of ongoing developments.

Despite the existence of numerous surveys, the literature remains dynamic and fragmented, with many reviews focusing on narrow sub-areas without providing a holistic perspective. Furthermore, critical differences in methodologies, performance metrics, and evaluation strategies are rarely juxtaposed explicitly, resulting in ambiguity and impeding direct comparison. In this context, our survey distinguishes itself by offering a comprehensive analysis and systematic contrast of major paradigms and frameworks within the field.

A significant point of contention among researchers lies in the ongoing debate regarding the primacy of model complexity versus data scale. Proponents of sophisticated architectures argue that novel inductive biases fundamentally transform generalization, while other camps emphasize the role of large annotated datasets and powerful optimization schemes. These and other philosophical and practical divergences form the foundation of heated discourse, as each approach harbors distinct advantages, limitations, and implications for scalability and real-world deployment. Our survey actively synthesizes these contrasting viewpoints, ensuring that contested domains and unresolved tensions are transparently presented for the benefit of both specialists and generalists.

To aid reader navigation and verification, all reference citations have been finalized for traceability (see Section **??**). For convenience, our detailed reference list, consolidated at the end of the survey, enables streamlined verification and follow-up.

Given the breadth of the field and the ongoing proliferation of techniques, this survey is organized to juxtapose, when appropriate, directly comparable methods in early summarizing tables. This structure empowers the reader to grasp major distinctions and key insights at a glance, without wading through dense technical expositions prematurely. We advocate that direct comparison and explicit contrast—especially where scholarly debate or divergence is active—forms the backbone of rigorous survey work.

To summarize, our contributions are: - Delivering a holistic, up-to-date synthesis spanning major AI subfields. - Highlighting and clarifying points of contention or divergence in current research. - Systematically contrasting differing methodologies to reinforce originality and reader clarity. - Providing finalized, traceable citations for full transparency.

Key insights from this survey, as well as open challenges and directions, are succinctly outlined at the end of each major section, facilitating targeted exploration by both novices and experts.

Table 1: Summary of Major AI Paradigms and Distinctive Features

Paradigm	Key Features	Typical Applications	Notable Challenges
Symbolic AI	Rule-based reasoning, interpretable logic	Expert systems, planning	Scaling, handling uncertainty
Machine Learning	Data-driven modeling, statistical learning	Classification, regression	Data quality, model interpretability
Deep Learning	Hierarchical representation learning, large neural nets	Vision, speech, language tasks	Data and computation intensity, explainability
Reinforcement Learning	Sequential decision-making, feedback-driven optimization	Game playing, robotics	Sample inefficiency, exploration-exploitation trade-off

1.1 Motivation for Automated, Data-Driven, and AI-Driven Analytics in Healthcare

The escalating complexity of modern healthcare, compounded by the proliferation of vast and heterogeneous data sources, has intensified the imperative for automated, data-driven analytic paradigms. Digital health data now originate from a wide array of modalities, including electronic health records (EHRs), imaging, wearable biosensors, and more. This diversity foregrounds the necessity of capabilities for real-time analysis and continuous monitoring, which are central to advancing personalized diagnostics, adaptive therapeutics, and overall operational efficiency.

Artificial intelligence (AI) and automation facilitate the extraction of actionable insights from high-velocity, complex data streams. These approaches have demonstrated efficacy across domains such as patient safety monitoring, disease surveillance, and optimization of clinical and administrative workflows. Nevertheless, despite these advances, the adoption of synthetic and AI-driven analytics continues to be hindered by several challenges:

Practitioner skepticism: Healthcare professionals remain cautious about entrusting critical decisions to algorithmic systems, citing concerns regarding reliability and interpretability.

Regulatory ambiguity: Established regulatory frameworks have yet to fully accommodate the nuances of automated decision-making in medicine.

High stakes of error: Unlike less critical sectors, errors or malfunctions in healthcare analytics can have profound, direct consequences for patient safety, underscoring the need for robust, transparent, and accountable computational strategies.

1.2 Emergence and Significance of Multimodal, Cross-Modal, and IoT-Enabled Platforms

Recent advances in healthcare analytics are increasingly propelled by the integration of multimodal and cross-modal data streams. These encompass imaging, genomics, unstructured clinical narratives, and physiological signals, collectively enabling the construction of comprehensive, patient-specific health models. The deployment of Internet of Things (IoT)-enabled platforms further enhances this landscape by supporting widespread, decentralized data acquisition through interconnected medical devices and in-home sensors.

Such systems are pivotal for continuous, context-aware health monitoring and the facilitation of remote care models. However, their utility is tightly coupled with the ability to harmonize and process data characterized by heterogeneous formats, distinct error characteristics, and diverse temporal resolutions [29, 68, 78]. Realizing this promise hinges on overcoming several technical challenges:

Real-time data fusion: Integrating diverse data streams in a temporally and contextually consistent manner.

Anomaly detection: Robust identification of clinically significant outliers or data errors across heterogeneous modalities.

Adaptive response: Enabling systems to dynamically adjust analytic or therapeutic strategies in response to incoming data.

Addressing these challenges remains a pressing research frontier that requires not only engineering ingenuity but also methodological innovation.

1.3 Central Themes: Big/Synthetic Data, Biomedical Signal Processing, and Real-Time Monitoring

Contemporary innovation in healthcare analytics is driven by paradigms that leverage both big data and synthetic data generation to enhance modeling, simulation, and decision support. Synthetic data, produced using techniques such as generative adversarial networks (GANs), variational autoencoders (VAEs), agent-based simulations, and natural language processing, serves several essential roles:

Policy prototyping enables testing of interventions or policies on simulated populations prior to real-world deployment. Privacy risk mitigation is addressed by reducing the chance of re-identification when sharing data for research and development. Data augmentation can supplement scarce or biased datasets, thereby improving model robustness and generalizability.

In parallel, advances in real-time biometric monitoring and biomedical signal processing are transforming proactive patient care and population health management. Yet, the reliance on synthetic and AI-generated data introduces new complexities, such as:

There are definitional ambiguities, with an ongoing lack of consensus regarding terminology and taxonomy for synthetic data categories. Model bias and auditability remain concerns, as there is an increased risk of embedding or exacerbating bias, accompanied by challenges in tracing model outputs to their data provenance. Privacy challenges persist, in part due to uncertainty regarding the adequacy of prevailing frameworks, such as HIPAA or GDPR, to regulate synthetic biomedical data [29]. For example, while differential privacy methods (such as PATE-GAN) may offer robust mathematical protection for synthetic datasets, their adoption in clinical contexts is limited by implementation complexity and a lack of standardization [29]. To ensure accountability, there are calls for frameworks such as digital chain-of-custody protocols encompassing encryption, regulated access, and possibly blockchain, to secure the integrity of synthetic healthcare data throughout its lifecycle [29]. Addressing these technical, ethical, and regulatory concerns remains fundamental for the responsible adoption of big and synthetic data in biomedical applications.

1.4 Scope and Organization of the Survey

This survey synthesizes contemporary research across the domains of multimodal data fusion, real-time biomedical signal processing, and IoT-enabled healthcare systems. Special attention is devoted to the technical, ethical, and regulatory complexities arising from the use of big and synthetic data within biomedical research and clinical practice [29, 68, 78]. The survey critically examines:

Methods for scalable analytics: Including computational frameworks capable of handling high-throughput, heterogeneous healthcare datasets. Privacy-preserving computation: Covering advances in differential privacy and secure, distributed computation, as well as innovations in digital chain-of-custody mechanisms. Secure data integration architectures: Focusing on distributed and federated approaches for robust, interoperable data management.

By integrating recent literature and application case studies, this work seeks to elucidate emerging best practices, articulate unresolved challenges, and recommend critical directions for the harmonization of data-driven, patient-centered innovation in the healthcare sector.

2 Foundations of Healthcare Analytics and Digital/IoT Infrastructure

The convergence of healthcare analytics with digital and IoT (Internet of Things) infrastructure has transformed patient care, hospital management, and medical research. Healthcare analytics leverages large-scale datasets generated from diverse sources—including electronic health records (EHR), sensors, mobile health platforms, and administrative systems—to support data-driven decision-making at every level. Central to these advances is robust digital infrastructure capable of securely collecting, storing, and transmitting vast amounts of sensitive clinical data.

The core of this infrastructure is the seamless integration of smart devices, cloud computing, and network protocols tailored to healthcare environments. This integration offers the potential for real-time patient monitoring, early disease detection, and personalized treatment regimens while presenting novel challenges in interoperability, data privacy, and system reliability.

A foundational tension in the development of healthcare analytics systems lies in balancing the needs for data accessibility and privacy. Approaches that emphasize open data sharing can accelerate discovery and innovation but risk exacerbating security and confidentiality concerns. Conversely, stringent restrictions on data flow may inhibit collaborative research and hinder the deployment of advanced analytics. There is also ongoing debate between proponents of centralized versus federated data architectures: centralized systems often enable comprehensive analytics but may constitute single points of failure, while federated approaches offer greater privacy but require sophisticated coordination among stakeholders.

Further, the adoption rate and practical effectiveness of these platforms differ by region and healthcare context, spurring ongoing discourse regarding standards, cost, and scalability. The landscape is rapidly evolving, with regulatory frameworks, technological capabilities, and clinical priorities shaping the trajectory of digital transformation.

In summary, the foundations of healthcare analytics and digital/IoT infrastructure represent both profound opportunities and considerable challenges. A nuanced understanding of the tensions between openness and privacy, as well as infrastructure design choices, is essential for realizing robust, ethical, and actionable healthcare analytics systems.

2.1 Evolution of Digital and IoT Health Systems

The digital transformation of healthcare infrastructure is fundamentally propelled by widespread digitization and the proliferation of Internet of Things (IoT) technologies, which collectively enable the development of highly connected health environments. The adoption of IoT-enabled health systems has facilitated the continuous collection and exchange of real-time patient data via a range of devices, including wearable sensors, mobile health applications, and smart diagnostic equipment. These systems support adaptive, context-aware electronic health services, promoting timely, individualized recommendations that transcend traditional clinical settings [10, 78]. The resulting distributed sensor landscape not only underpins robust health monitoring but also enables the deployment of algorithmically guided interventions, as digital platforms leverage real-time analytics to personalize care and improve the management of chronic diseases [78].

The implementation of digital health technologies—such as health information technology systems, electronic health records (EHRs), and interoperable IoT device frameworks—has significantly expanded both the scope and scalability of healthcare analytics. Advances in federated and cloud-based architectures have facilitated large-scale integration of diverse health data sources, enhancing operational efficiency and supporting advanced clinical research [8, 10, 78, 94]. Contemporary infrastructures enable real-time data ingestion in conjunction with high-throughput analytics. Large healthcare institutions, therefore, can automatically aggregate data streams from clinical, radiological, laboratory, and patient-generated sources, providing the foundation for predictive analytics, surveillance, and personalized medicine initiatives [78, 94].

Nonetheless, the expansion of these systems has exposed persistent inefficiencies, particularly concerning health information exchange, workflow interoperability, and the sustainable deployment of AI-powered applications. These challenges are especially pronounced in rural or under-resourced environments, where resource disparities hinder the effective realization of digital health benefits [10, 94].

A critical examination reveals that the scalability and utility of digital health systems are frequently constrained by non-technical factors. Policy frameworks, for example, exert a major influence on the operational viability and legal boundaries of digital health integration by affecting adoption rates, promoting or hindering interoperability among vendors and platforms, and shaping approaches to patient data governance [8, 10]. Despite technical readiness at the device level, issues related to the lack of standardized protocols, inconsistencies in data representation, and proprietary barriers regularly impede seamless cross-platform data exchange. Ethical considerations also gain prominence as digitization accelerates, raising urgent issues pertaining to privacy, algorithmic transparency, and the digital divide. The demand for rigorous governance models that prioritize data standardization, promote explainability in algorithmic decision-making, and ensure transparent patient consent

Table 2: Key Dimensions of Healthcare Analytics and Digital/IoT Infrastructure

Dimension	Description	Opportunities	Challenges
Data Sources	EHRs, sensors, mobile health, admin systems	Rich, multimodal insights	Data heterogeneity, integration difficulties
Digital Infrastructure	Smart devices, cloud, network protocols	Scalability, real-time monitoring	Interoperability, system reliability
Data Architecture	Centralized vs. Federated systems	Comprehensive analytics vs. privacy preservation	Single point of failure, coordination complexity
Security & Privacy	Access controls, encryption, compliance	Trust, regulatory alignment	Confidentiality risks, innovation obstacles
Adoption & Standards	Regional, organizational variations	Tailored implementation, global reach	Scalability, cost-effectiveness

Table 3: Comparative Overview of Digital Health System Evolution: Benefits and Persistent Barriers

Aspect	Key Advancements	Exemplar Systems/Statistics	Ongoing Challenges
Digitization & Adoption	Sharp increase in EHR and health IT uptake, e-prescribing, and patient access to records [10]	10-fold increase in EHR usage (hospitals); 85%+ e-reporting of results; 97% hospitals with online records access [10]	Uneven access in rural/small providers; technical complexity; digital divide
Integration & Interoperability	Multisource, real-time data integration; emergence of hospital-wide big data platforms [94]	WCH-BDP: data from 27+ systems, 20 PB storage, automated/self-service retrieval [94]	Data non-standardization; workflow interoperability; proprietary barriers
Analytics & AI Readiness	Large-scale analytics platforms underpin predictive, personalized care and AI deployment [8, 78]	AI-powered detection (e.g., pulmonary nodule at 98.8% accuracy); ML/DL study accuracies up to 98.5% [8, 94]	Handling streaming heterogeneous data; scalability; missing data; quality control [8, 78]
Governance & Ethics	Expansion of patient engagement and data-driven care models, with federated and cloud architectures [10, 94]	Increased transparency, metadata and terminology standards, 5S security in hospital platforms [94]	Privacy concerns; fragmented regulations; inconsistent patient consent; algorithmic transparency [10, 94]

is at odds with the current landscape characterized by variable implementation quality and fragmented regulatory environments across jurisdictions [8, 10, 78, 94].

Active debates continue regarding the prioritization of interoperability standards versus local/regional flexibility, the balance between strict privacy controls and data availability for research, and the extent to which AI-powered analytics can be responsibly integrated given concerns around explainability and bias [8]. Divergence exists regarding optimal deployment strategies in varied contexts: while developed nations focus on advanced analytics, resource-constrained and rural environments struggle with basic digital access and sustainability [10, 94]. These ongoing discussions underscore the field's rapidly evolving and sometimes contentious nature, indicating significant opportunities for innovation in governance, technical, and ethical domains.

Summary: In summary, the evolution of digital and IoT health systems has yielded substantial advancements in interoperability, real-time analytics, and individualized care, yet it is also accompanied by unresolved challenges such as fragmented standardization, varying data quality, privacy concerns, and the digital divide between diverse healthcare environments. Addressing these issues remains essential for realizing the full potential of digital health infrastructures [8, 10, 78, 94].

2.2 Core Concepts, Data Modalities, and Taxonomies

Objectives and Scope. This subsection aims to (i) define foundational concepts central to multimodal analytics in healthcare; (ii) provide a structured taxonomy of key data modalities and their analytical challenges; (iii) critically synthesize technical and practical trade-offs of data fusion strategies, emphasizing real-world evidence; and (iv) directly link modality challenges with regulatory and governance frameworks. These objectives ensure comprehensive alignment with leading survey best practices and facilitate downstream discussion of technical and societal implications.

At the heart of modern healthcare analytics lies the integration of multiple data modalities, which requires a systematic conceptual framework for effective cross-modal data fusion, representation, and interpretation. Key concepts include:
Cross-modal learning: Leveraging relationships between distinct data types for improved inference and knowledge transfer.

Multimodal representation: Encoding and integrating heterogeneous sources—such as imaging, genomics, biosignals, and clinical text—into unified computational models [8, 11, 15, 34, 43, 44, 49, 67, 78, 83, 91, 93, 101].
Multi-view fusion: Aligning complementary perspectives (e.g., EHR text and diagnostic images) to enhance predictive accuracy and model robustness.
IoT sensor integration: Incorporating continuous real-time data capture from sensor streams, thereby promoting context awareness and enabling active feedback mechanisms for both clinical and home settings [8, 49, 67, 78].

Healthcare data modalities are both diverse and complex, necessitating clear taxonomies for rigorous analysis. We provide a comprehensive categorization below, which directly informs the analytic and regulatory challenges faced by current cross-modal systems:
Omics data: Genomics, transcriptomics, proteomics, and metabolomics, which characterize molecular profiles.
Medical imaging: Including CT, MRI, PET, and ultrasound, foundational for diagnostic and prognostic modeling.
Biosignals: Such as EEG, ECG, and sEMG, offering physiological and functional insights.
Clinical data: Both structured (e.g., lab results) and unstructured (e.g., EHR notes).
Behavioral, audio, and video data: Supporting remote assessment and patient engagement.
IoT-derived data streams: High-frequency, continuous sensor data enabling granular monitoring and timely intervention [8, 27, 28, 35, 47, 51, 53, 61, 69, 77, 78, 87, 102].

The integration of these modalities underscores the increasing analytical complexity of healthcare, with each domain presenting distinct representational, computational, and governance challenges. Notably, data standardization, patient privacy, cross-institutional interoperability, and regulatory compliance (e.g., GDPR, national regulations) are critical themes that deeply intertwine with technical development, as highlighted in recent policy studies [10, 34, 94].

Persistent Technical and Regulatory Challenges. Despite significant advances, persistent heterogeneity in data formats, sampling rates, and annotation standards continues to hinder large-scale multimodal analytics. Multi-institutional and multi-vendor

datasets are often sparse, fragmented, or marked by errors and anomalies [8, 10, 24, 78, 94]. This limits the generalizability and fairness of derived models, and raises consequential regulatory and ethical concerns over data validity, patient safety, and transparency. Furthermore, weak or non-linear correlations between modalities (e.g., loosely coupled sensor streams and clinical events) challenge simplistic fusion strategies and drive the need for sophisticated alignment, imputation, and representation learning methods. These limitations are worsened during real-world deployments, where missing data, class imbalance, and noisy streams are prevalent [24, 78, 94]. Such settings amplify the importance of robust data governance and traceability standards as called for in recent global policy initiatives [10, 34, 94].

Recent developments in cross-modal learning—such as cross-attention mechanisms, graph neural networks with modality-aware encoding, and prompt-based large language models—reflect concerted efforts to close these modality-driven gaps [15, 34, 43, 49, 67, 83, 91]. Nevertheless, key obstacles remain, including the challenge of achieving robust representational alignment under domain shifts, the opacity of current fusion methods, and ongoing concerns about robustness to missing or low-quality modalities.

A central unresolved methodological debate is the choice between early, intermediate, and late fusion, as well as the tension between modality-specific versus joint representation learning [34, 44, 93]. Early fusion allows joint learning from raw modalities but suffers acutely from sample misalignment and scale disparities. Late fusion enables modular, interpretable post-hoc integration but can limit synergistic effects—a recurring concern in studies evaluating imaging versus molecular data integration [8, 24, 53, 83]. For instance, in molecular oncology and radiomics, the lack of robust, unified benchmarks underscores ongoing controversies as to which approach generalizes best in cross-site, real-world settings [35, 53, 83]. The interpretability of fusion mechanisms (often criticized as "black box") is increasingly seen as a prerequisite for both regulatory and clinical acceptance [34, 93].

Furthermore, comparative studies have illuminated both the advantages and shortcomings of rival methodologies or standards tailored for distinct domains and tasks (e.g., radiology versus behavioral analytics), with substantial divergences observed in generalizability, explainability, and resilience to dataset shift [8, 24, 53, 83]. These issues are tied directly to the evolving regulatory landscape, as new global AI health partnerships and policy recommendations advocate for open benchmarks, interoperable data standards, and traceable, reproducible analytic pipelines [10, 34, 94].

To provide a structured overview and direct, comparative reference for readers, Table 4 summarizes central data modalities, core analytical challenges, and illustrative applications in multimodal healthcare analytics. This table is tightly linked to the analytic and governance narrative above, highlighting both technical barriers and their implications for practical adoption, privacy, and regulatory compliance.

Table 4 highlights the intersection of technical and governance hurdles for each modality, reinforcing this subsection's core message: technical, regulatory, and ethical challenges in multimodal

healthcare analytics are inextricably linked and require joint solutions.

To address these challenges, emerging taxonomies and systematic benchmarks of multimodal fusion strategies—as well as landmark datasets and policy initiatives from the past 2–3 years—are proving essential to driving practical advances in healthcare analytics, supporting efforts from both technical and regulatory communities [8, 10, 34, 43, 49, 94].

Summary and Takeaways. In summary, this subsection has defined foundational principles; mapped core data modalities, taxonomies, and technical barriers; and underscored the necessity of embedding cross-modal analytics within interoperable, transparent, and ethically compliant frameworks. Drawing direct connections between technical, regulatory, and real-world considerations, it sets an explicit analytic baseline for the deeper methodological and governance debates that follow in subsequent sections.

2.3 Datasets, Benchmarks, and Standards

Objectives: This subsection aims to (1) catalog key multimodal healthcare datasets, (2) critically evaluate the evolution and challenges of benchmarking practices across digital and IoT health analytics, and (3) synthesize current debates on data standards, harmonization, and their intersection with regulatory and practical considerations.

The empirical foundation for digital and IoT health analytics is provided by high-quality, multimodal datasets and robust benchmarking repositories. Several prominent datasets—such as Uni-Mod1K, ImageNet-ESC, LPBA40, IXI, OASIS, ADNI, BraTS, CheXpert, and the MIT-BIH Arrhythmia Database—encompass diverse data types including imaging, biosignals, and longitudinal clinical records, all accompanied by established annotation protocols for algorithmic benchmarking [3, 8, 17, 19, 38, 48, 55, 58, 61, 77–79, 83, 86, 96, 102]. The integration of real-time IoT analytics platforms enables the assessment of algorithms on heterogeneous and high-velocity streaming data, a capability integral to emerging paradigms of remote and continuous care [78, 83].

Table 5 below offers a synthesized overview of representative multimodal healthcare datasets, directly linking each to primary modalities, data formats, and core use cases. This table serves as an essential bridge between the descriptive listing above and the analysis that follows, illustrating the diversity and breadth of resources driving both foundational research and application-oriented benchmarking in the field.

As shown in Table 5, the scope of available datasets reflects both the breadth of multimodal input types and the evolving targets for benchmarking in healthcare AI. Several key trends emerge: the rise of cross-modal datasets for few-shot and transfer learning [58, 102], the increasing integration of diverse imaging sequences and clinical records [19, 79], and the value of expert-annotated, high-resolution biosignal datasets for tasks such as ECG classification [86]. These trends directly inform algorithmic developments, as seen in advances ranging from transformer-based image registration [19] to deep learning models for X-ray and MRI classification [53].

Despite these advances, the process of dataset selection and benchmarking remains highly contested. Major methodological controversy centers on the standardization of annotation protocols

Table 4: Summary comparison of principal healthcare data modalities, typical analytical challenges, and illustrative applications in multimodal learning.

Modality	Common Analytical Challenges	Representative Applications	Relevant References
Omics data	High dimensionality; batch effects; sparse annotation; data privacy	Disease risk prediction; pharmacogenomics	[8, 47]
Medical imaging	Heterogeneous formats; volumetric/temporal variation; annotation scarcity; reproducibility	Diagnosis, segmentation, radiomics	[27, 35, 53, 61, 69]
Biosignals	Noise and artifacts; temporal misalignment; class imbalance; interoperability	Seizure detection; cardiac monitoring	[49, 67, 77]
Clinical data	Unstructured text; missing/corrupted values; privacy/security; semantic interoperability	Patient outcome prediction; cohort analysis	[8, 10, 24, 78, 94]
Behavioral/audio/video	Limited labeled data; inter/intra-subject variability; privacy, consent, regulatory compliance	Remote screening; symptom tracking	[8, 11, 77, 102]
IoT-derived streams	High frequency; continuity gaps; device heterogeneity; data governance	Real-time monitoring and alerting; context-aware	[8, 49, 78]

Table 5: Summary of Representative Multimodal Healthcare Datasets and Supported Modalities

Dataset	Primary Modalities	Data Type(s)	Use Case(s)
UniMod1K	Imaging, Biosignals, Language	Multimodal (RGB, Depth, Text)	Multimodal learning, object tracking [102]
ImageNet-ESC	Imaging, Audio	Images, Audio files	Cross-modal few-shot learning, benchmarking [58]
LPBA40	MRI	3D Images	Brain segmentation, neuroimaging [19]
IXI	MRI	3D Images	Neuroimaging, brain mapping [19]
OASIS	MRI, Clinical	Images, EHR	Alzheimer's research, diagnosis [19]
ADNI	MRI, PET, Clinical	Images, EHR, Cognitive Scores	Alzheimer's disease progression, federated learning [79]
BraTS	MRI	Multisequence MRI	Brain tumor segmentation, federated benchmarking [79]
CheXpert	X-ray	Images, Labels	Chest disease classification, deep learning [53]
MIT-BIH Arrhythmia	ECG	Biosignals	Arrhythmia detection, ECG classification [86]

and benchmarking criteria. In radiomics and imaging, initiatives such as the Image Biomarker Standardisation Initiative (IBSI) are crucial for harmonizing feature definitions, processing pipelines, and validation metrics in order to control for variability and improve reproducibility [1]. The technical literature, including comprehensive reviews [35, 84], documents persistent heterogeneity in acquisition protocols, segmentation methods, and feature nomenclature. For instance, whereas the IBSI enables comparability across CT, PET, and MRI modalities, broad implementation is hindered by site-specific workflow differences and resistance to fully prescriptive methodologies—particularly in advanced PET/CT applications, where the biological meaning of extracted features is debated and the clinical impact of standardized versus locally adaptive algorithms is still unresolved [84].

The arrival of multimodal and streaming IoT data has amplified these tensions, with new benchmarks demanding fresh performance metrics that move beyond static accuracy. Recent clustering around metrics like response ratio, variation factor, and real-time accuracy under variable connectivity [78] demonstrates both the need for context-appropriate benchmarks and the risk of metric fragmentation. Works on cloud-based medical data analytics highlight active debate over prioritizing absolute accuracy versus service responsiveness and reliability within resource-constrained, heterogeneous IoT environments [8, 78]. Real-world failures—such as degraded responsiveness under unstable connectivity or reduction in generalizability due to non-standardized streaming formats—underscore that consensus on optimal benchmarking is still nascent, and ongoing.

Harmonization efforts face both organizational and technical barriers. Many datasets originate from single-center studies or demographically narrow populations, limiting generalizability and increasing the risk of bias in downstream models [38, 58, 78]. The widespread reliance on proprietary IoT devices further fragments

the ecosystem, leading to divergent approaches toward federated benchmarking and privacy preservation [78, 79, 94]. For example, federated learning methods have seen increasing adoption to bypass regulatory barriers and privacy issues in medical imaging, but they introduce new challenges of data heterogeneity, validation, and secure aggregations [79].

Progress toward open-source, multi-institutional datasets continues, exemplified by resources such as new lung ultrasound repositories with thorough expert annotation and standardized metadata [48]. These initiatives are reinforced by the development of cross-domain validation frameworks in genomics and medical imaging [42, 47], which bring a degree of procedural rigor but also raise substantive questions about balancing universal standards with the need for local adaptation to clinical workflow or resource constraints.

Critical Analysis and Unresolved Debates: The ongoing push for interoperability and common standards—driven by national initiatives and global partnerships [10, 24, 94]—directly impacts both methodological innovation and regulatory/ethical frameworks. However, practical deficiencies remain: persistent incompatibilities, the lack of widely accepted standards across modalities, and continued controversy over the trade-off between strict universal benchmarks and context-specific customization [24, 94]. Methodological debates linking benchmarks to analytic strategies (for example, the merits of early versus late data fusion) remain sharply divided, especially where rival approaches may yield different levels of robustness or interpretability in heterogeneous, real-world scenarios.

Summary: In sum, standardized datasets and benchmarks are essential for the credible evaluation and trustworthy translation of digital and IoT healthcare analytics. However, entrenched technical,

organizational, and regulatory complexities—intertwined with unresolved methodological controversies—demand sustained, multidisciplinary collaboration among domain experts, standards bodies, and technology stakeholders. Prudent navigation of these challenges will determine the pace and equity of progress in AI-driven healthcare innovation.

2.4 Data Privacy, Security, and Governance

Subsection Objective: This subsection aims to explicitly define and review the measurable objectives, fundamental requirements, and open challenges in securing patient privacy, ensuring data integrity, and instituting effective data governance within digital and IoT-enabled health platforms. It seeks to directly map the intersection of technical, regulatory, and ethical frameworks to enable trustworthy, equitable, and operationally robust healthcare analytics.

The exponential scale-up of digital and IoT health systems has brought urgency to safeguarding patient privacy, guaranteeing data security, and establishing comprehensive data governance frameworks. Electronic Health Records (EHRs) and IoT health platforms exhibit unique vulnerabilities—such as data breaches, unauthorized access, and failures in transparency—that can erode patient trust and disproportionately impact marginalized communities, challenging the ethical legitimacy of digital health programs [10, 11, 24, 34, 35, 40, 44, 45, 47, 51, 52, 60, 61, 73, 78, 79, 89, 90, 93, 94, 98]. These risks are compounded by systemic disparities in technology access and digital infrastructure, which exacerbate inequities in the secure use of digital health solutions [78, 94].

Transitions to regulatory and ethical frameworks are essential to anchor technical innovations within enforceable policies. National and supranational frameworks—including HIPAA in the United States and GDPR in the European Union—set minimum obligations for privacy, stewardship, and patient rights. However, as IoT-enabled and AI-driven systems rapidly evolve in scale and capability, these policies face ongoing tension and reinterpretation regarding applicability and enforcement [6, 8, 10, 11, 22, 24, 26–28, 30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 79, 89, 90, 92–95, 98–100, 102]. For example, IoT healthcare platforms generate persistent data via dynamic device connectivity and consumer-grade sensors, necessitating specialized technical strategies such as federated learning and on-device analytics, as well as robust data minimization—yet these measures must be implemented without loss of analytical utility [8, 24, 94]. Regulatory frameworks, however, often struggle to keep pace with these technological advances, creating implementation gaps and ambiguous responsibilities among stakeholders.

To comprehensively address these challenges, overlapping technical safeguards and procedural innovations must be embedded within robust governance models. Fine-grained, role-based access controls and real-time auditing, together with formalized consent management, algorithmic transparency, and clear documentation, form the backbone of privacy-centric, ethical-by-design systems focused on patient autonomy and societal accountability.

The interplay of these considerations is particularly acute for smaller or under-resourced providers, who often lack the resources for consistent, effective compliance [10, 78, 94]. As seen in leading hospitals and platforms [10, 94], emerging best practices—such as

automated data retrieval, metadata governance, multi-modal data integration, and transparent auditing—offer operational blueprints, yet translation of high-level regulatory intent to daily reality remains a significant hurdle. These enforcement gaps can inadvertently impede innovation or widen disparities in health access and outcomes [10, 78, 94].

Open Challenges and Research Gaps: Despite progress, digital and IoT health systems face unresolved and interconnected issues, requiring integrated responses at both the technical and regulatory levels: Regulatory-Technology Gap: Regulatory cycles lag rapid innovation, especially with IoT, multi-modal fusion, and AI-driven analytics [8, 24, 94]. Scalable, Equitable Governance: Existing governance structures and technical safeguards lack scalability for resource-limited settings, perpetuating access inequities [10, 78, 94]. Standardization and Interoperability: Fragmented, unstandardized, and legacy data infrastructures impede secure, seamless health information exchange [10, 24, 79, 94]. Methodological Transparency: Opacity in algorithmic processes and reporting undermines oversight and patient trust, an issue intensifying with the proliferation of AI and ML in sensitive domains [24, 34, 79, 94]. Data Minimization and Utility: Striking a balance between strict privacy and the need for high-utility analytics remains an active research challenge as modality diversity and data volume expand [8, 24]. Operationalizing Consent and Autonomy: Consent management remains procedural, with enforcement proving challenging at scale, especially across dynamic and longitudinal IoT data environments [94]. By explicitly linking these technical and procedural challenges to their respective regulatory and governance implications, organizations and policymakers can better develop scalable and actionable solutions.

In summary, the foundational components of digital and IoT health analytics—including system infrastructures, multi-modal data integration, standardization, benchmarking resources, and privacy-centric governance—must undergo continual co-evolution with regulatory and ethical frameworks. Improved transitions between technical and legal discussions clarify that achieving trustworthy and impactful healthcare analytics depends on aligning these safeguards and innovations. Future directions should prioritize developing scalable, equitable governance models, interoperable standards, explicit algorithmic accountability, and operational consent mechanisms—ensuring that technical progress is matched by social benefit and regulatory compliance.

Subsection Summary: To build trustworthy, effective, and equitable digital health platforms, both technical and procedural safeguards must be fully integrated with evolving regulatory and ethical frameworks. Strengthening the linkage and operationalization between technical advances and governance is crucial for overcoming persistent gaps in privacy, security, and access. Explicit, actionable objectives—such as interoperable infrastructures, transparent analytics, scalable consent management, and equitable data stewardship—should guide future research and policy to ensure that digital health innovation leads to measurable improvements in health outcomes and social trust.

2.5 Biomedical Signal Processing and Real-Time Health Monitoring

This subsection begins by defining its measurable objectives: (1) to systematically synthesize methodological innovations in biomedical signal processing and real-time health monitoring, (2) to deeply analyze technical and regulatory challenges, and (3) to critically compare competing approaches in terms of accuracy, efficiency, interpretability, and privacy. These aims are pursued with an emphasis on how advanced computational and learning techniques drive accurate, real-time physiological data interpretation, as well as how emerging policies and global initiatives shape practical deployments.

Recent progress in biomedical signal processing is largely attributed to the proliferation of wearable and IoT-enabled sensors, which now facilitate ongoing clinical and at-home monitoring. Core technical challenges include enhancing biosignal acquisition under ambulatory conditions, developing robust artifact removal and denoising algorithms, pioneering lightweight yet accurate learning techniques for edge devices, and supporting timely detection of adverse health events. Continuous advances in these domains are increasingly shaped by new public datasets, benchmark releases, and landmark AI health partnerships over the past two to three years, which have contributed both technical resources and frameworks for regulatory alignment.

Researchers leverage diverse datasets—spanning ECG, EEG, and PPG signals—for real-time monitoring of vital signs and anomaly detection. A wide array of signal processing and machine learning frameworks has been proposed. Deep learning models, for example, offer high pattern recognition capacity but often depend on large datasets and substantial computational resources, posing obstacles for resource-limited embedded systems. By contrast, classical signal processing and shallow learning methods are more lightweight and interpretable but may not generalize as well to complex multimodal or temporal tasks. Notably, debates surrounding analytic strategies—such as early (feature-level) versus late (decision-level) data fusion—remain unresolved. Early fusion can exploit richer cross-signal dependencies but risks propagating noise, while late fusion offers modularity and resilience but may underutilize complex interactions. Real-world failures and case studies highlight that improper fusion strategy selection can undermine detection accuracy or inflate system latency, especially in heterogeneous patient populations.

Comparative analysis reveals sharp trade-offs among accuracy, latency, computational cost, and interpretability. Although certain methods reach state-of-the-art performance in controlled environments, they may struggle to generalize to diverse global populations or prove robust against sensor displacement and motion artifacts. Landmark initiatives in recent years have released more demographically-inclusive datasets and promoted open evaluation challenges, catalyzing new standards for robustness and fairness.

Transitioning from technical to regulatory and ethical considerations, it is clear that advancing biomedical signal processing necessitates careful alignment with data privacy regulations, security frameworks, and guidelines for the responsible use of personal health data. Open problems straddle benchmarks, privacy, and analytics: the absence of standardized evaluation datasets frustrates

fair comparisons; securing sensitive patient data, especially under streaming or federated learning paradigms, remains a rapid-growth research area; and technical advances in multimodal integration must be balanced with practical constraints such as battery life and real-time response.

Open Challenges and Research Gaps. A number of persistent challenges remain. First is the need for widely adopted, standardized evaluation datasets and benchmarks to ensure transparent, reproducible algorithm comparison. Second, there is ongoing difficulty in achieving robust generalization across heterogeneous patient groups and diverse physical contexts. Third, privacy-preserving processing and data security are critical issues—solutions must not only comply with regulations but also maintain analytic utility. Lastly, integrating multimodal inputs in resource- and battery-constrained systems while preserving responsiveness is a central area of research. Approaches such as adaptive learning and edge/federated processing offer promise, though scaling these solutions for longitudinal and global monitoring remains difficult.

In summary, biomedical signal processing for real-time health monitoring is an interdisciplinary, rapidly evolving field. Ongoing work converges toward improving robustness, computational efficiency, interpretability, and privacy protection, with growing influence from global policy initiatives and regulatory benchmarks. Addressing methodological, analytic, and ethical challenges in a cohesive manner is essential for the realization of reliable, pervasive, and equitable health monitoring technologies.

2.5.1 Signal Interpretation and Disease-Specific Applications. **Objectives and Scope:** This subsection aims to survey state-of-the-art techniques for biomedical signal interpretation and their application in disease-specific contexts. We explicitly focus on methods advancing real-time health monitoring and intelligent decision-making for neurological, cardiac, and prosthetic applications using multimodal signals such as EEG, sEMG, and ECG. The section discusses leading techniques, benchmarking practices, and outstanding challenges, highlighting unique frameworks and contemporary paradigms relevant to translational digital health.

Biomedical signal processing forms the foundation of precise, real-time health monitoring, facilitating sophisticated interpretation of physiological signals such as EEG, sEMG, and ECG. With the advent of advanced quantitative methodologies, disease-specific applications have achieved significant advancements. Notably, intraoperative EEG analysis employing the Hurst exponent as a principal feature has yielded an objective, quantitative means to identify transitions between anaesthetic states, thus enhancing assessment accuracy and providing a critical alternative to subjective clinical judgement during surgery [71]. Such methodological innovations exemplify how the integration of time-series dynamics extends clinical utility.

In the realm of prosthetics, surface electromyogram (sEMG) signals—despite their nonlinear character and susceptibility to noise—have been effectively leveraged for intuitive prosthetic control. Recent developments employ multiresolution decomposition via dual-polynomial interpolation, optimizing denoising and the reconstruction of motor-evoked signals. This approach facilitates

reliable multi-class motion decoding within noisy real-world environments, effectively translating complex biosignals into responsive, user-adaptive prosthetic commands [80]. The robustness of these preprocessing pipelines is central to the real-time applicability of prosthetic systems.

The computational analysis of ECG signals further illustrates the transformative impact of advanced processing strategies on disease detection. Transforming one-dimensional ECG traces into two-dimensional time-frequency representations—particularly using the continuous wavelet transform (CWT)—has measurably improved arrhythmia classification accuracy. A recent comparative study systematically evaluated both CWT and short-time Fourier transform (STFT) methods for transforming ECG segments from the MIT-BIH Arrhythmia Database, finding that CWT-based representations, especially those using the Ricker wavelet, generally outperformed STFT for arrhythmia classification tasks. Using transfer learning with pre-trained convolutional neural network architectures such as ResNet-18, the study achieved a top binary classification accuracy of 96.17% for normal versus premature ventricular contraction (PVC) beats. Although STFT yielded faster processing per segment, CWT provided superior localization of abnormal beats as demonstrated by Grad-CAM visualizations, enhancing interpretability for clinical review. The study also highlighted the significant role of hyperparameter tuning in transformation methods and confirmed the practicality of fine-tuned pre-trained networks in real-time cardiac monitoring systems. Notably, limitations of this work included a focus on binary classification and a single dataset, indicating future research should address multiclass ECG problems, custom model designs, improved visualization techniques, and validation on additional datasets [86].

Reliable benchmarking is underpinned by widely recognized datasets, including MIT-BIH for ECG, NinaPro for sEMG-based prosthetic studies, and CHB-MIT as well as other EEG repositories for neurological disorder monitoring [71, 80, 86]. These curated datasets serve as reference standards, bolstering reproducibility and standardization essential for rigorous comparative algorithmic development and translational research.

Despite these advances, significant challenges persist in the real-time interpretation of multimodal signals, particularly in practical settings involving data streams from wearable and IoT-enabled platforms. The integration of heterogeneous data sources, each with varying sample rates, modalities, and transmission reliability, complicates timely and accurate monitoring—especially during acute clinical scenarios [2, 37, 49, 78]. Clinical applications such as seizure and home-based epilepsy monitoring are emblematic of these challenges, introducing artifacts, motion-induced noise, and missing data that necessitate advanced artifact rejection and robust adaptation strategies.

Innovative multimodal systems exemplify the cutting edge of real-time monitoring. Advanced seizure detection and prediction platforms have achieved robust results by leveraging multiresolution analysis (e.g., maximal overlap discrete wavelet transforms, MODWT) to decompose EEG signals and extract informative frequency-band features. As demonstrated in recent work, patient-specific multiresolution convolutional neural networks (CNNs) process frequency-band decomposed EEG frames and aggregate consecutive frame outputs within a sliding window to improve sensitivity

and minimize false positive rates for preictal state classification. On public benchmark datasets such as CHB-MIT and Kaggle/American Epilepsy Society intracranial EEG, these approaches achieved sensitivities of up to 85% with correspondingly low false prediction rates, outperforming prior methods in several scenarios [37]. For additional context, Table 6 summarizes comparative sensitivity and false prediction rates for leading seizure prediction approaches from [37].

Furthermore, multimodal pipelines that integrate wearable EEG with inertial measurement units (IMUs)—enabling concurrent analysis of brain activity and movement—have demonstrated reliable reduction of false alarms in home environments. Specifically, pipelines utilizing EEG-IMU fusion with targeted artifact rejection through movement sensor data and multi-source training deliver improved sensitivity and substantial reductions in false detections, outperforming previous automated and even expert neurologist-reviewed approaches on at-home seizure detection tasks. The integration of movement artifact handling and cluster-based under-sampling for class imbalance has been pivotal in these advancements, with future work planned for broader validation and deep learning extensions as more data become available [18]. Techniques such as advanced multi-view deep feature learning for EEG further enhance interpretability and performance by combining traditional frequency and time domain information with deep feature representations, providing explainable outputs via interpretable machine learning classifiers [2].

Identified Gaps and Future Directions: Key challenges remain in integrating heterogeneous, multimodal data sources in resource-constrained or highly variable conditions, developing robust and generalizable artifact rejection strategies, and improving the interpretability of deep neural features, particularly for patient-specific applications. Future research should prioritize solutions for multiclass classification (e.g., for ECG), custom model architectures, richer visualization and explainability of decision processes (especially in CNN-based pipelines), and adaptations that account for patient-specific and location-dependent signal variability. In addition, rigorous transferability and cross-dataset validation are essential for translating results to broader clinical and home settings. Enhanced collaboration between domain researchers and clinical practitioners will be critical for closing gaps between algorithmic innovation and practical healthcare delivery.

The synergy between multimodal feature extraction and intelligent artifact rejection supports continuous monitoring beyond traditional clinical settings and embodies a pivotal stride toward pervasive, patient-centric digital health.

2.5.2   Advanced Feature Learning and Biometric Monitoring Algorithms. This subsection surveys recent advances in feature learning paradigms and algorithmic frameworks developed to improve the efficacy and interpretability of biometric monitoring systems. The objective is to elucidate the role of deep and hybrid architectures in real-time biomedical signal processing, highlight unique frameworks such as HRLS-TS and AMV-DFL, and discuss ongoing challenges along with directions for future research.

Table 6: Benchmark Sensitivity (%) and False Prediction Rate (FPR) for Seizure Prediction Approaches

Dataset	Sensitivity (%)	FPR
CHB-MIT	82	0.058
Kaggle	85	0.19
State-of-art (CHB-MIT, 13 pts)	90.2	0.0713

The progression of feature learning paradigms has markedly improved both the accuracy and interpretability of biometric monitoring systems. Deep and hybrid architectures—including recurrent neural networks (RNNs) and long short-term memory networks (LSTMs)—successfully capture temporal dependencies inherent in biomedical time series. Building on these, frameworks such as the Hybrid Recurrent Long Short-term based Tyrannosaurus Search (HRLS-TS) algorithm have emerged, enabling real-time health monitoring with adaptive signal discrimination in dynamic conditions [49]. The integration of bio-inspired metaheuristics, exemplified by the Tyrannosaurus Search optimizer, further enhances algorithmic performance by balancing predictive accuracy with computational efficiency.

A significant advancement in the field is the adoption of interpretable, multi-view feature integration. The Advanced Multi-View Deep Feature Learning (AMV-DFL) framework represents this evolution by combining conventional frequency-domain features (e.g., extracted via fast Fourier transform) and time-domain features from raw signals with automatically-learned deep features sourced from one-dimensional convolutional neural networks (1D CNNs) [2]. This multi-view approach synthesizes a richer, more comprehensive feature representation that has demonstrated superior effectiveness over single-view or traditional methods, notably improving detection accuracy in biomedical signal analysis such as EEG-based seizure detection. Within this framework, interpretable machine learning classifiers, including multi-view forests, are employed in tandem with explainability mechanisms such as tree-based SHAP (T-XAI), which provide quantifiable insights into feature importance and channel contributions. Specifically, the AMV-DFL framework outperforms models using only traditional or single-view deep features by 4% and other state-of-the-art methods by an average of 3% in accuracy, thereby aiding clinicians in identifying EEG features relevant to epileptic seizures and potentially discovering novel biomarkers [2]. These advances directly support clinical decision-making and help address the regulatory demand for transparency in AI-driven monitoring solutions [2, 49].

Despite these advances, several core challenges and open research directions remain. Achieving reliable real-time signal optimization is critical, particularly as wearable and IoT-based health monitoring platforms typically face stringent power and computational constraints. Future research should focus on developing adaptive strategies that further balance model sophistication with energy efficiency and minimal computational latency. Additionally, there is a need for robust, interpretable frameworks that can scale to multimodal signal analysis and facilitate clinical translation by supporting continuous, personalized, and auditable monitoring in contemporary healthcare settings [49]. Cross-disciplinary benchmarking and validation of these frameworks across diverse patient populations and signal types also represent essential next steps.

2.5.3 Evaluation and Real-World Validation. **Objectives and Scope:** This section surveys contemporary evaluation practices for biomedical signal processing systems within large-scale, real-time health monitoring frameworks. Emphasis is placed on operational performance metrics, real-world validation strategies, and persistent challenges, with attention to recent advances documented across multicenter deployments, hospital data platforms, and wearable device trials.

The successful translation of biomedical signal processing research into large-scale, real-world health monitoring depends on robust evaluation protocols. As IoT-enabled healthcare grows, reliance on traditional accuracy metrics alone has become insufficient. To address operational realities, a range of performance indicators—including service response ratio, delivery time, variation factor, identification ratio, and aggregate processing time—are regularly employed to assess system viability under diverse streaming data and network conditions. These metrics permit comprehensive appraisal of a platform's responsiveness and resilience in the face of fluctuating loads, real-world device failures, and transient data disruptions [78].

The adoption of such metrics (Table 7) allows for critical insight into the operational strengths and limitations of signal monitoring frameworks.

A central priority in contemporary validation efforts is the utilization of diverse, multicenter, and authentic clinical datasets. Data amalgamated from large-scale hospital infrastructures and wearable device deployments ensures sufficient heterogeneity for comprehensive benchmarking, exposing analytic algorithms to an array of physiological, demographic, and environmental confounders [8, 18, 26, 49, 54, 77, 94]. For example, the WCH-BDP platform provides an integrated, high-volume clinical data environment that supports robust, real-time analytics and paves the way for scalable AI deployments in hospital settings [94], while recent multicenter trials employing wearable EEG and movement sensors have advanced the use of multimodal data streams for artifact rejection and real-world seizure detection [18]. Cardiac arrest prediction pipelines leveraging large, event-rich datasets [54] and machine learning approaches for diverse conditions such as autism and chronic disease [8, 77] further highlight the need for external validation and the development of transferable, generalizable models. This strategy strengthens model generalizability and facilitates identification of persistent challenges, including missing data, context-dependent adaptation, and site-specific bias—issues increasingly addressed through federated and transfer learning approaches.

Table 7: Representative Metrics for Evaluating Real-Time Health Monitoring Systems

Metric	Description
Service Response Ratio	Measures the fraction of health events correctly recognized and responded to within a defined period
Delivery Time	Quantifies time taken from signal acquisition to response generation
Variation Factor	Assesses system performance stability under changing network or signal conditions
Identification Ratio	Evaluates correct identification of target events relative to all system outputs
Aggregate Processing Time	Captures total computational time for processing and classification tasks

Artifact rejection and adaptive recalibration remain vital to reliable deployment. Notably, multicenter evaluations in seizure detection demonstrate that the combination of auxiliary sensing modalities (e.g., IMUs) and advanced postprocessing significantly reduces false alarm frequency without compromising sensitivity [18]. The growing integration of explainable AI techniques further enhances clinician trust, providing actionable insights and facilitating safe and effective intervention.

Gaps and Future Challenges: Key open challenges include ensuring model robustness to missing or low-quality data, developing frameworks that support privacy-preserving, multicenter data integration, and sustaining performance across heterogeneous, real-world operating conditions [8, 94]. The scalability of analytics in distributed hospital environments [94], reliable artifact handling in pervasive sensing [18], and dynamic adaptation to context and user-specific variability represent important directions for future work, with ongoing research focused on federated and transfer learning, adaptive recalibration, and interpretable AI.

In summary, the discipline is transitioning from incremental algorithmic improvements to the deployment of comprehensive, integrated pipelines that emphasize signal fidelity, interpretability, computational efficiency, and operational robustness—collectively supporting the broader adoption and real-world impact of real-time health monitoring technologies.

3 Cross-Modal, Multimodal, and IoT-Driven Healthcare Analytics

This section explicitly aims to survey recent advances, frameworks, and challenges in cross-modal and multimodal analytics within IoT-driven healthcare settings. The primary objectives are to delineate the scope and unique role of multimodal data integration for healthcare, identify key approaches and paradigms, and critically analyze unaddressed gaps and future research opportunities. The survey scope includes methodologies enabling the fusion of heterogeneous health data (such as medical imaging, physiological signals, and electronic health records) and their application within the context of Internet of Things (IoT) infrastructures for improved diagnostic, predictive, and personalized healthcare.

We review representative frameworks that demonstrate cross-modal integration strategies, highlighting their operational paradigms that enable synergy across diverse data streams. Emphasis is placed on novel architectures and learning mechanisms that support real-time analytics and decision support within resource-constrained IoT environments, as well as the trade-offs between accuracy, efficiency, and system scalability.

Several persistent gaps and challenges remain in this field. Notably, current multimodal systems face limitations in harmonizing noisy, incomplete, or temporally misaligned data modalities, which can undermine their diagnostic reliability. Privacy-preserving mechanisms and robust security protocols specifically tailored for multi-source health data in distributed IoT deployments also lack maturation. Furthermore, there is a need for explainable AI paradigms that provide transparency into cross-modal inferencing—a critical requirement for clinical adoption.

Concrete future research directions include the development of adaptive fusion strategies that dynamically handle missing or asynchronous modalities, optimization of secure data sharing protocols within federated IoT networks, and the establishment of benchmarks for evaluating cross-modal methods with clinically meaningful metrics. There is also substantial scope to explore transfer learning and domain adaptation techniques, enabling improved generalization across diverse patient populations and healthcare settings.

By analyzing these elements, this section seeks to position itself as a comprehensive and original reference for researchers and practitioners interested in the intersection of multimodal analytics and IoT-driven healthcare innovation.

3.1 Proportionate Data Analytics (PDA) and Data Management

The rapid expansion of Internet of Things (IoT) devices within healthcare ecosystems has culminated in the generation of voluminous, heterogeneous, and high-velocity data streams. These streams frequently exhibit variable formats, fluctuating quality, and diverse error profiles—characteristics that challenge traditional analytic frameworks, which implicitly assume uniform reliability among data sources. Such conventional models prove inadequate in environments where data quality is nonstationary and prone to both transient anomalies and persistent sensor failures.

Proportionate Data Analytics (PDA) offers a principled paradigm for addressing these complexities. By employing statistical techniques—for example, linear regression applied to temporally disjoint intervals—PDA frameworks can discriminate between routine signal variability and substantive anomalies, thus enabling the rapid identification and isolation of compromised data streams without excessively impacting system responsiveness or service continuity [78]. This context-sensitive approach allows healthcare systems to dynamically tailor data processing pipelines: streams exhibiting anomalous behavior, incomplete information, or degraded quality are proportionately de-emphasized or routed for secondary

verification, while reliable signals maintain their operational priority. In practice, this strategy enhances overall analytic robustness and ensures that clinical decision-support systems remain well-calibrated, even in the face of environmental noise and transient uncertainty.

The imperative for PDA in healthcare stems from two principal motivations:

- **Mitigating Downstream Analytic Risk:** To ensure that decision-support mechanisms are governed by the most reliable data, thereby minimizing the possibility of error amplification and unintended clinical consequences.
- **Autonomous Adaptive Resource Allocation:** To empower monitoring platforms and autonomous diagnostic systems with the capacity to allocate analytic resources dynamically, guided by real-time assessments of data quality and provenance.

Crucially, PDA frameworks embed anomaly detection, quality scoring, and error flagging directly into the data ingestion and management layers, thereby maximizing operational transparency. By systematically documenting points of error, correction, and exclusion throughout the analytic workflow, PDA not only serves technical robustness but also aligns with regulatory mandates for traceable data lineage and auditability in clinical contexts.

3.2 Multimodal and Multisource Data Fusion

Subsection Objectives: This subsection provides a focused overview of how multimodal and multisource data fusion is shaping healthcare analytics. It restates the broader survey goal of mapping the landscape of integrative AI methodologies in healthcare and highlights the specific aims to (1) clarify core data fusion paradigms, (2) comparatively synthesize the benefits and limitations of each, and (3) identify pressing open challenges and future research needs in this domain.

The confluence and synthesis of diverse biosignals, medical images, behavioral metrics, and IoT-derived contextual data have revolutionized the scope and depth of healthcare analytics. Integrative multimodal and multisource fusion methodologies are indispensable for leveraging the distinctive yet complementary informational content inherent to each modality, whether in continuous physiological monitoring, static or temporal imaging, longitudinal behavioral profiling, or the incorporation of structured electronic health records [8, 27, 28, 30, 31, 33–35, 42, 47, 49, 51, 53, 56, 61, 77, 78, 87, 94]. Key application areas include:

Radiomics. This encompasses the extraction and fusion of high-dimensional features from CT, MRI, or PET imaging with omic datasets and clinical histories, thereby supporting enhanced biomarker discovery and risk stratification [35, 53, 56, 61]. Despite the promise, radiomic approaches face open challenges in reproducibility, big data management, integration of multimodal and temporal data, and explaining models for clinical trust.

Cardiometabolic Monitoring. Integrating wearable biosensors with behavioral and environmental data enables robust, longitudinal risk assessment [8, 31, 77]. Nevertheless, systems must address missingness and variability in real-world data streams, as well as the need for models generalized to diverse populations.

Intelligent Hospital Platforms. Cohesive synthesis of laboratory results, procedural logs, and real-time monitoring supports advanced analytics, AI-driven triage, and operational optimization [78, 94]. While real-time data integration systems have shown scalable impact, further work is needed for structural standardization, privacy governance, and multicenter interoperability.

To effectively harness multimodal data, several distinctive fusion paradigms have emerged, summarized below:

Systems frequently require dynamic selection among these paradigms, dictated by the evolving data landscape, operational constraints, and analytic objectives (see Table 8). Each fusion approach has context-specific strengths and limitations. For instance, early fusion readily models short-range interactions but is challenging with high-dimensional or partially missing data, while late fusion is flexible but may sacrifice synergistic interpretation; hybrid methods strike a balance but increase computational complexity and design choices [8, 26, 30, 87, 94].

A central technical challenge in this context remains the quantification and propagation of uncertainty, which is vital for trustworthy diagnostic inference and risk management in clinical settings [25, 34, 35, 61]. Methods such as Dempster-Shafer theory have been adapted to healthcare data fusion to aggregate evidential support, capturing both the strength of belief and the degree of conflict among modalities [25]. Probabilistic and evidential frameworks facilitate a nuanced response to incomplete or contradictory data. Technical strategies for handling missingness and partial observation include:

Domain-Informed Statistical Imputation. Leveraging clinical expertise enables plausible data reconstruction but may be limited by evolving datasets and domain drift.

Representation Learning with Variable-Length Sequences. Models robust to interrupted or incomplete input streams are emerging as critical, especially in IoT-intensive and event-driven contexts [6, 8, 26, 77, 94].

Comparative Synthesis. In addition to the above, fusion methods differ in their computational requirements, explainability, sensitivity to heterogeneous data quality, and ability to leverage structured versus unstructured inputs; the choice of paradigm is application-dependent and drives downstream effectiveness [34, 53, 61, 94].

Open Challenges and Future Directions. *(1) Robustness and Generalizability:* Improving resilience to missing, noisy, or inconsistent modalities remains an urgent research gap, as highlighted by persistent limitations in real-world medical image and biosignal analysis [8, 35, 53, 56]. *(2) Standardization and Interoperability:* There is a pronounced need for data curation, harmonization, quality assurance, and common protocols to support reliable fusion, especially in multi-institutional and cross-national contexts [35, 42, 47, 51, 94]. *(3) Explainability and Transparency:* Ensuring trustworthy, explainable fusion models is a key open challenge identified throughout the literature [8, 28, 34, 35, 53], particularly as AI tools are integrated into decision-making pipelines. *(4) Computational Scalability:* Efficient real-time analytics for massive, heterogeneous multimodal data streams highlight unresolved issues in distributed and edge-computing scenarios [78, 94]. *(5) Ethical and Privacy Considerations:* Multisource fusion heightens ethical complexity, reinforcing the critical need for privacy protection, secure governance, and responsible AI deployment [34, 94].

Table 8: Comparison of Multimodal Data Fusion Paradigms

Paradigm	Characteristics	Advantages and Limitations
Early Fusion	Merges raw features from all modalities before modeling; enables low-level cross-modal interactions	Captures direct synergies; prone to dimensionality issues, may require large datasets
Late Fusion	Combines outputs from modality-specific models (decision-level)	Robust to missing modalities; may overlook deep synergistic structure
Joint/Hybrid Fusion	Interleaves intermediate representations, often with attention mechanisms or layered interactions	Balances representation richness with tractable modeling; excels in streaming and heterogeneous IoT environments

Collectively, these innovations and open questions constitute the methodological backbone and forward research agenda for developing highly resilient, data-rich analytics in real-world healthcare systems. This focus is essential to fulfill the overarching survey objective of charting trustworthy, effective, and scalable AI/Machine Learning integration across multimodal medical contexts.

3.3 Emerging Adaptive and Cross-Modal Processing

This subsection aims to synthesize and critically assess recent advances in adaptive and cross-modal learning paradigms for healthcare-related data, with an explicit focus on adaptive, distributed, and semi-supervised strategies that address challenges unique to multi-institutional, privacy-sensitive, and weakly supervised data scenarios. By tightly connecting these objectives to the broader aims set forth in the abstract and introduction, this section specifically targets readers interested in the intersection of robust machine learning, multi-modal data integration, and practical deployment for next-generation healthcare environments.

The dynamic, often multi-institutional, and privacy-sensitive character of contemporary healthcare data ecosystems has driven the advancement of learning paradigms that are adaptive, distributed, and effective under limited annotation. In distributed IoT healthcare, adaptive, semi-supervised, and federated multi-source fusion strategies have demonstrated substantial promise: enabling analytics at scale even when labeled data is scarce, privacy regulations are restrictive, and data distributions drift over time [49, 63, 78]. Approaches like Proportionate Data Analytics (PDA) [78] illustrate the importance of segregating errors and adapting analytics to varying stream characteristics, which is critical in global and low-resource healthcare contexts where data heterogeneity is high and service reliability must be maintained.

Self-supervised learning facilitates the extraction of generalized data representations from large unlabeled corpora, thus supporting efficient transfer to downstream clinical tasks. Continual learning methods aim to prevent catastrophic forgetting by permitting analytic models to adapt to new data streams; this property is essential for long-term monitoring and responding to the evolution of disease profiles. Semi-supervised cross-modal frameworks such as SPamCo [63] employ co-training, regularization, and pseudo-labeling across multiple data views. Notably, SPamCo allows distributed parallelization and rigorous PAC learnability guarantees, enhancing scalability and robustness—a particular advantage for large-scale clinical or sensor data in resource-constrained environments. These advancements are directly relevant for healthcare settings characterized by pervasive sparsity and noise in labeled data, as well as substantial underlying heterogeneity.

Nevertheless, important methodological limitations and research gaps remain. Current adaptive fusion models still demonstrate vulnerability to performance declines under extreme distributional shifts produced by temporal changes, contextual factors, or institutional heterogeneity within incoming clinical or IoT data [78]. While federated learning mitigates privacy and localization barriers, it poses new problems—such as difficulty in aligning learned representations across non-i.i.d. datasets, as well as challenges in ensuring the consistency and stability of distributed model updates. Methods like the Robust Cross-modal Learning (RCL) framework [36] highlight the importance of handling partially mismatched pairs in cross-modal retrieval scenarios by emphasizing negative pair information and risk minimization, demonstrating robust performance under weak supervision and high label noise. However, the operational integration of these techniques under dynamic, real-world healthcare conditions—particularly for global or low-resource deployments—remains an unresolved challenge.

Furthermore, debate exists regarding the best mechanisms for harmonizing adaptive, privacy-preserving, and cross-modal learning within unified frameworks. Several approaches address individual aspects (e.g., error segregation in IoT data streams [78], self-paced co-training [63], negative-pair focused cross-modal retrieval [36]), yet a fully synthesized perspective that tightly couples robust uncertainty quantification, scalable fusion, and reliable adaptation is still underexplored relative to prior surveys. Notably, the survey advances the discourse by explicitly integrating recent advances such as SPamCo's multi-view self-paced co-training, RCL's robust treatment of noisy multimodal data, and PDA's dynamic analytics for IoT health, thereby charting a more comprehensive path for resilient multimodal healthcare analytics.

A persistent point of controversy is the extent to which optimism about cross-modal fusion and adaptation should be tempered by candid recognition of the brittleness of current techniques to out-of-domain inputs and practical deployment barriers. For instance, while methods such as RCL [36] and SPamCo [63] show robustness under specific experiment settings, their generalizability under real-world constraints and limited resources is often less certain and merits careful further study.

Balancing these strengths and weaknesses, the unique contribution of this section—relative to the overall survey and to earlier reviews—lies in its critical synthesis of robust cross-modal retrieval (as through RCL [36]), practical distributed data handling for real-time IoT healthcare (notably PDA [78]), and the latest advances in adaptive, self-paced co-training and federated learning (SPamCo [63]), while situating these methods within the urgent context of privacy, global applicability, and healthcare system transformation.

For future research, key directions include: (1) developing adaptive fusion schemes resilient to severe distributional drift and out-of-distribution data, (2) constructing privacy-aware cross-modal architectures that ensure consistency across federated updates and heterogeneous sources in both well-resourced and low-resource health settings, and (3) advancing unified theoretical and practical frameworks that bridge self-supervised, semi-supervised, and

continual learning for healthcare's diverse modalities. Achieving progress in these areas is a prerequisite for the transformational impact of multimodal and IoT-driven analytics in next-generation healthcare systems [36, 49, 63, 78].

In summary, this subsection has critically reviewed the objectives, strengths, and current limitations of adaptive and cross-modal learning paradigms, articulated unique integrative opportunities that go beyond earlier reviews, provided counterpoints and highlighted ongoing controversies, and delineated the section's distinct contributions within the survey's overall goals. These syntheses and perspectives aim to guide the continuing evolution of robust, scalable, and privacy-preserving AI in healthcare and to inform a broad readership spanning researchers, practitioners, and policymakers.

4 Machine Learning, Deep Learning, and Explainable AI in Healthcare

This section aims to systematically review the applications, progress, and challenges of machine learning (ML), deep learning (DL), and explainable AI (XAI) within the healthcare domain. Our objectives are to critically survey major methodologies, synthesize distinguishing frameworks proposed in the field, highlight contrasts with prior works, and outline future research priorities, in direct alignment with the overarching aims established in the abstract and introduction. We specifically intend to provide comparative insights into recent advances and synthesize unresolved questions, targeting researchers, clinicians, and practitioners seeking a comprehensive understanding of current capabilities and persistent gaps.

Machine learning and deep learning have catalyzed significant advancements in healthcare, providing robust capabilities for large-scale data analysis, predictive analytics, and personalized medicine. Recent years have witnessed a growing deployment of ML/DL approaches for diagnosis, prognosis, and treatment recommendation; however, widespread clinical integration remains hindered by several methodological and practical limitations. A major barrier continues to be the opacity of many state-of-the-art deep models, resulting in limited trust and interpretability for end users, a subject of ongoing debate in both the AI and clinical communities.

Explainable AI (XAI) methodologies seek to bridge the interpretability gap, augmenting transparency for clinicians and stakeholders. While various explanation frameworks and post-hoc interpretability techniques have emerged, their adoption is neither universal nor undisputed. Many approaches are narrowly tailored to specific tasks or data modalities, and a consensus regarding standardized evaluation metrics for explainability in clinical settings remains elusive, with some proposing user-centric measures but with limited agreement or implementation.

Despite considerable progress, several critical research gaps persist and spark ongoing controversy. First, the majority of clinical ML/DL models are developed and assessed on limited or homogeneous datasets, raising concerns about generalizability, particularly in low-resource and global health contexts where data scarcity and heterogeneity are acute. Second, bias and fairness issues, often stemming from non-representative data or unaccounted confounders, are seldom systematically addressed despite their central

importance in real-world applications. Third, the predominant paradigm still favors accuracy-centric evaluation over comprehensive assessments of safety, robustness, and real-world impact, an imbalance which can obscure vulnerabilities of AI systems under clinical deployment. Fourth, while this survey synthesizes distinct frameworks for integrating XAI into clinical workflows, there remains a pressing need for universal benchmarks and guidelines to harmonize evaluation of interpretability methodologies across diverse healthcare domains and populations.

In contrast to previous reviews, this survey makes the unique contribution of emphasizing the intersection of ML/DL methodologies with context-aware explainable frameworks, especially in multimodal and real-world healthcare scenarios. We explicitly compare the strengths and limitations of popular model architectures and explanation algorithms across heterogeneous clinical applications, including underrepresented and low-resource settings. By doing so, we develop and present comparative criteria that facilitate a more rigorous and nuanced assessment of clinical AI solutions beyond traditional performance metrics, highlighting both established and emerging evaluation criteria.

Furthermore, to enhance coherence with the survey's overall goals, we reiterate our focus on fostering unified understanding and practical adoption. We aim not only to synthesize the state-of-the-art but also to foreground open controversies and identify directions for resolving persistent obstacles.

To further advance the state of the art, future research should prioritize the development and validation of models with improved transparency, generalizability, and fairness, particularly in resource-limited and diverse clinical environments. Additional priorities include the creation of standardized benchmarks and reporting guidelines for clinical XAI; deeper investigation into the integration of end-user feedback within explanation frameworks; and rigorous evaluation of methods across real-world clinical workflows and geographies. The reconciliation of optimism about transformative impact with candid discussion of methodological pitfalls is critical to responsible and equitable advancement.

In summary, this section provides a critical and balanced overview of ML, DL, and XAI in healthcare, explicitly highlighting both transformative potential and prevailing challenges—while reinforcing this section's contribution to the survey as a whole. By foregrounding open research questions, acknowledging ongoing controversies, and emphasizing concrete future directions, we aim to clarify and unify priorities for researchers and clinicians seeking to responsibly deploy AI-driven solutions in healthcare.

4.1 Model Architectures and Learning Methods

4.1.1 Traditional ML and Deep Learning Architectures in Healthcare.
The landscape of machine learning (ML) and deep learning (DL) architectures in healthcare has evolved considerably, shaped by the increasing diversity and complexity of biomedical data, alongside escalating demands for diagnostic accuracy, scalability, and robustness. Traditional ML algorithms—specifically, Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Decision Trees—have long provided robust baselines for structured data analyses. These algorithms possess notable strengths in settings emphasizing interpretability and computational efficiency, often

yielding competitive performance in classification and regression tasks where datasets are moderate in size and efficient feature engineering is feasible [44, 91].

The advent of high-dimensional and heterogeneous biomedical data modalities—including medical imaging, multi-channel physiological signals, and structured electronic health records (EHRs)—has made deep neural architectures indispensable. Among these, Convolutional Neural Networks (CNNs) have become the predominant choice for medical image analysis, excelling at complex feature extraction and hierarchical representation learning in imaging modalities such as X-ray, MRI, and CT [3, 8, 22, 33, 34, 38, 41, 44, 49, 61, 69]. Furthermore, architecture variants such as U-Net and 3D CNNs are tailored for biomedical image segmentation and volumetric delineation, effectively handling spatial complexities inherent in such data [55, 56, 59, 61].

For sequential clinical data, particularly physiological time-series (e.g., ECG, EEG) and EHR data, Recurrent Neural Networks (RNNs)—notably Long Short-Term Memory (LSTM) networks—have demonstrated substantial efficacy in capturing temporal dependencies [58, 69, 87, 101]. The emergence of Transformer architectures has further augmented the modeling of long-range temporal and contextual dependencies in both time-series and textual modalities. The incorporation of multi-head self-attention within Transformers enables robust performance and flexibility, particularly in complex multimodal fusion scenarios [6, 8, 19, 26, 28, 41, 44, 53, 55, 102].

Emerging models such as Graph Neural Networks (GNNs) are increasingly leveraged to analyze data with intrinsic graph structures, enabling significant advancements in tasks ranging from molecular property prediction in drug discovery to the modeling of population or knowledge networks in epidemiology [76, 92, 102]. Integrative frameworks combining GNNs and large language models (LLMs) within multi-modal architectures have shown marked improvements in property prediction and knowledge extraction across complex scientific domains [92].

Despite these advances, several persistent challenges limit the full realization of deep learning's promise in healthcare: Data annotation scarcity, impeding the training of data-hungry models High inter-class variability, which complicates generalization Limited generalizability to rare or outlier cases Insufficient transparency and interpretability, hindering clinical acceptance These obstacles remain focal points for ongoing research and methodological innovation [28, 30, 33, 53, 55, 56, 59, 93].

4.1.2   Transfer, Hybrid/Ensemble, Annotation-Efficient, and Self-Supervised Learning for Multimodal and IoT Healthcare Data. To address the persistent challenges of data scarcity, annotation cost, and heterogeneous data sources in healthcare AI, a range of advanced learning paradigms—encompassing transfer learning, hybrid/ensemble methods, annotation-efficient approaches, and self-/semi-supervised learning—have been developed. These approaches collectively improve model utility for both unimodal and multimodal healthcare datasets by increasing data efficiency, robustness, and generalizability.

Transfer learning leverages pre-trained models, often developed on large and general or related datasets, which are then fine-tuned on more specific biomedical or clinical datasets. This allows researchers to capitalize on existing large-scale knowledge and

adapt it for more specialized applications, proving especially valuable when labeled data is limited or expensive to obtain [6, 8, 28, 41, 53, 69, 92]. For example, frameworks such as COMET utilize RNNs pre-trained on extensive EHR cohorts, adopting both early and late fusion strategies to integrate omics and clinical data. Such approaches not only enhance predictive accuracy, particularly for small or rare-sample scenarios, but also provide deeper insights into biological mechanisms [55].

Hybrid and ensemble methods, including model stacking and the fusion of traditional machine learning (ML) classifiers with modern deep learning (DL) architectures, are crucial for improving generalizability and mitigating overfitting—critical aspects in high-stakes fields such as disease diagnosis for oral cancer, cardiovascular conditions, and rare event prediction [8, 27, 33, 53, 56, 61, 91]. These strategies facilitate dynamic adaptation to different contexts and requirements, striking a balance between interpretability and predictive performance.

Annotation-efficient paradigms, notably semi-supervised and self-supervised learning frameworks, are essential under constraints of limited annotated data. Semi-supervised learning techniques, such as pseudo-labeling and consistency regularization, enable models to harness large quantities of unlabeled data. In tasks like prostate MRI segmentation, these methods can attain results closely approaching fully supervised models [6, 26, 47].

Self-supervised learning designs pretext or contrastive tasks that allow models to learn powerful representations from unlabeled data alone. These approaches are particularly prominent in transformer-based and contrastive learning settings and have demonstrated notable success in the extraction of discriminative biomedical features [6, 30, 51, 69, 78]. In complex environments such as multi-sensor IoT healthcare, self-supervised models—often enhanced with knowledge distillation or robust pseudo-annotation pipelines—achieve significant gains in annotation efficiency and performance [6, 8, 26, 28, 53, 55, 78, 79].

Multimodal learning, which integrates distinct types of information such as text, images, structured records, and physiological waveforms, is increasingly promising for health informatics. Multimodal large language models (MLLMs), equipped with sophisticated fusion and alignment components—such as joint attention mechanisms—enable cohesive contextual analysis and improved clinical prediction [6, 8, 17, 28, 41, 47, 53, 55, 58, 61, 78]. Augmenting these architectures with transfer learning and self-supervised strategies further strengthens their capacity to generalize to complex real-world settings. Persistent challenges remain, however, most notably in the areas of model interpretability, computational requirements, and the need for well-aligned multimodal representations.

The approaches summarized in Table 9 collectively represent significant advancements for overcoming key hurdles in data efficiency, generalizability, and multimodal information fusion within healthcare AI.

4.2   Explainability, Transparency, and Clinical Trust

4.2.1   The Imperative for Explainable AI (XAI) Methods and Clinical Applications. Realizing the transformative potential of ML/DL in healthcare critically depends not only on predictive accuracy,

Table 9: Overview of Learning Methods and Their Key Contributions in Healthcare

Method	Main Contribution	Key Applications/Examples
Transfer Learning	Leverages pre-trained models to improve performance, especially with limited labeled data	Fine-tuning CNNs for medical imaging [8, 41, 53], RNNs on EHRs (COMET) [55]
Hybrid/Ensemble	Combines diverse models to boost generalizability and reduce overfitting	Disease diagnosis (oral cancer, CVD) [8, 33, 56]
Semi-supervised Learning	Utilizes both labeled and unlabeled data, improving learning efficiency	MRI segmentation [26, 47], pseudo-labeling [6]
Self-supervised Learning	Learns representations via pretext tasks, requiring no labeled data	Transformer-based feature extraction [30, 69, 78]
Multimodal Learning	Fuses diverse data types for richer predictions and context integration	MLLMs with joint attention for clinical text, images, and waveforms [6, 41, 55, 78]

but equally on model explainability, transparency, and the establishment of clinical trust. The need for explainable AI (XAI) is particularly acute in medical decision-making, where clinicians and regulatory stakeholders demand that machine learning models deliver reliable outputs accompanied by meaningful justification for decisions [28, 32, 54, 61, 82].

Post-hoc XAI techniques—such as SHAP, LIME, and ELI5—are now standard tools to quantify feature contributions and elucidate model decision pathways [28, 44, 61, 64, 82]. These techniques have found concrete application in recent diagnostic models: for diseases like diabetes, monkeypox, and cardiac arrest, XAI methods highlight the influence of clinical or physiological features, enabling patient-level interpretability. For instance, diagnostic work on monkeypox employed SHAP, LIME, and ELI5 to explicitly identify fever, oral/genital ulcers, swollen lymph nodes, and fatigue as the most influential predictors in an interpretable risk model. This transparency directly supports clinical adoption, bridges the gap between advanced AI and healthcare practitioners, and fosters trust with real-world impact [82]. Similarly, in time-series models for cardiac arrest prediction in ICU settings, the application of SHAP revealed which ECG-derived heart rate variability measures—such as the baseline width of the RR interval histogram—contributed most to risk estimation, turning deep model predictions into actionable insights for clinicians [54]. Tools like these are thus foundational in linking algorithmic outputs with clinical reasoning and facilitating integration into practice, while also supporting regulatory and operational transparency [2, 28, 32, 44, 61, 62, 64, 82].

Cutting-edge areas of XAI are gaining prominence in clinical research and practice. Event-level causal reasoning is one such domain, where models utilizing causal interventions are able to uncover and separate spurious correlations among features or modalities, leading to more robust and clinically relevant insights about temporal and causal relationships in diagnostic decisions [32, 62, 64]. Another key area is the interpretability of signal-level data, including biomedical signals such as EEG and ECG: post-hoc approaches (SHAP, counterfactuals, event-level causal modeling) not only differentiate relevant temporal features but also contextualize explanations using established physiological phenomena. Cardiac arrhythmia detection, for example, leverages XAI tools to attribute predictions to signal features like R-R interval regularity, while advanced deep feature learning frameworks for EEG utilize posthoc explainability methods to connect learned representations with known electroclinical markers, thereby supporting not only accurate prediction but also the discovery of new potential biomarkers [2, 39, 67]. Despite these advancements, a significant challenge remains in mapping deep model features to explicit physiological or pathological entities, emphasizing the ongoing need for refinement. The dynamic landscape of XAI in healthcare thus requires

continuous development to ensure that data-driven intelligence aligns with the realities of clinical knowledge, decision-making, and workflow [2, 39].

4.2.2 User-Centric Analytic Interfaces and Transparency in IoT Healthcare. Recent developments in explainable analytics for IoT healthcare increasingly emphasize user-centric interfaces that extend beyond mere algorithmic transparency to enable real-time, patient- and case-specific interpretability, facilitating direct clinician and patient involvement. Self-explanatory analytic platforms are now designed to integrate predictive modeling with intuitive explanation layers, with approaches such as SHAP providing granular, transparent insights tailored for both expert and non-expert users [20, 39, 77, 78].

In the context of IoT-enhanced healthcare, system transparency coupled with the capability for rapid, real-time analytics is essential for meeting regulatory requirements and supporting clinical adoption [8, 78]. The implementation of transparent analytic interfaces and dynamic visualization tools fosters user trust, clinical acceptance, and ongoing engagement. These features enable continuous anomaly detection, instant feedback, and system adaptability in response to dynamic IoT data streams and evolving clinical situations [20, 39]. For example, interfaces that embed SHAP-based explanations allow both clinicians and patients to inspect the reasoning behind diagnostic or predictive outputs, thereby increasing confidence and supporting evidence-based decisions [20, 39]. Furthermore, proportionate data analytics systems exemplify how differentiating between data variations and errors in heterogeneous streams—by employing dynamic classification and responsive feedback mechanisms—can bolster real-time performance, reliability, and compliance in complex IoT healthcare networks [78]. Collectively, these advanced explainability and transparency approaches lay a robust foundation for user-engaged, trustworthy analytics integral to continuous, patient-centered care in IoT-enabled healthcare systems.

4.2.3 Clinical and Real-World Validation of AI and IoT-Augmented Systems. The definitive measure of ML/DL and XAI systems' value lies in their rigorous clinical and real-world validation. This process demands continual evaluation across heterogeneous patient populations, diverse clinical workflows, and variable operating environments, with explicit objectives such as measurable improvements in clinical outcomes (e.g., reduced event rates or improved diagnostic accuracy), as well as quantifiable performance metrics like area under the receiver-operating characteristic curve (AUROC), weighted mean absolute percentage error (WMAPE), or accuracy. For instance, multicenter studies evaluating AI-augmented radiological segmentation or real-time risk monitoring consistently report that models integrating explainability and user-centric design are

more readily adopted and trusted in clinical settings due to their transparent and clinically relevant outputs [2, 26, 39, 49, 54, 94].

Specific examples highlight the operational impact of these systems. Lee et al. [54] demonstrated that their real-time cardiac arrest prediction model achieved an AUROC of 0.881 and an AUPRC of 0.104 in ICU environments, outperforming traditional clinical parameter models in a retrospective cohort of over 5,000 ICU stays, and underscoring the importance of evaluation metrics tied directly to patient safety outcomes. Wang et al. [94] describe a hospital-wide big data platform enabling AI-augmented applications such as pulmonary nodule detection with 98.8% accuracy at scale, which increased clinical data retrieval efficiency and research productivity. In structural health monitoring for healthcare infrastructures, Feng et al. [26] introduced a data fusion method that maintained WMAPE as low as 1.57%–15%, even under sparse or noisy sensor conditions, exemplifying robust generalization and actionable decision support in real-world scenarios. Furthermore, Ahmad et al. [2] achieved 3–4% accuracy improvements in epileptic seizure detection by leveraging multi-view deep feature learning, illustrating the translation of enhanced model performance into clinically relevant biometric event detection.

Despite these successes, robust generalizability remains a pervasive challenge, especially when models confront shifts in data distributions, patient demographics, imaging and scanning protocols, or rare and emergent disease presentations. Addressing this, best practices include external validation with independent datasets to ensure transferability across sites; prospective and multicenter studies to capture broader patient and workflow variability; and continuous adaptation or refinement using real-world user feedback. These strategies are essential for sustaining long-term model robustness and for ensuring persistent clinical relevance [26, 39, 49].

Bridging detailed technical advances with operational and clinical significance, it is evident that fusion strategies and transfer learning approaches directly support scalable deployment. For example, the integration of big data governance in clinical systems [94] and data fusion in heterogeneous monitoring [26] have demonstrably facilitated cross-department clinical decision making and streamlined research-to-practice translation.

In summary, the synergy of advanced learning algorithms, efficient annotation and transfer learning strategies, and the imperative for explainability and transparency encapsulates both the promise and enduring challenges for AI-driven healthcare. A principled integration of these components, guided by measurable clinical and technical outcomes, is essential to realize trustworthy, effective, and equitable machine intelligence in clinical practice.

5 Medical Imaging, Multimodal, and Cross-Modal Analytics

Section Objectives and Scope

This section aims to systematically review and analyze the latest advances in medical imaging analytics, with a particular focus on multimodal and cross-modal data integration strategies. The objectives are: (1) to delineate key technical frameworks for multimodal data fusion and cross-modal knowledge transfer, (2) to propose a conceptual taxonomy for integration strategies, (3) to elucidate operational trade-offs and targeted healthcare outcome metrics

for these approaches, and (4) to bridge the gap between technical innovation and clinical scalability via real-world deployment examples.

Conceptual Framework: Taxonomy for Multimodal Integration

To provide structural clarity to discussions in this section, we introduce a taxonomy of multimodal and cross-modal analytics in medical imaging. Integration strategies are classified along two principal axes: (1) *Fusion Depth* (early, intermediate, late), describing at what stage data from different modalities are combined, and (2) *Interaction Type* (direct fusion, shared latent space, cross-modal distillation), distinguishing explicit feature combination from implicit alignment or transfer. This framework facilitates structured comparison across studies and highlights operational trade-offs.

Summary and Structure

The following subsections specify technical approaches and measurable outcomes for each major integration strategy. Throughout, we interleave bridging commentary to connect architectural details to their implications for clinical scaling, such as robustness to data variation, deployment feasibility, and impact on patient outcomes.

Such a structured analysis spotlights the dual challenges of algorithmic sophistication and translational viability, setting the stage for detailed technical exposition and reflective discussion on real-world adoption.

5.1 Automated Segmentation, Registration, and Imaging Diagnostics

This section aims to synthesize recent advances and ongoing challenges in the field of automated medical image analysis, focusing on segmentation, registration, and diagnostic support. Our primary objective is to distill state-of-the-art methodologies, highlight limitations of current paradigms, and articulate open research questions for future exploration within neuroimaging, oncology, and cardiology applications.

Advances in automated medical image analysis have substantially enhanced both the scale and accuracy of diagnostic workflows across domains such as neuroimaging, oncology, and cardiology. Central to this progress are deep learning architectures—including DenseNet and 3D-UNet—that demonstrate state-of-the-art performance in complex segmentation tasks by effectively capturing hierarchical and contextual features from high-dimensional imaging data. For example, a recent approach utilizes a two-stage network architecture, employing a 2D DenseNet for filtering followed by a 3D-UNet with integrated dense and residual blocks, to achieve robust coronary artery segmentation from CCTA images. This setup not only yields high Dice Similarity Coefficients but also streamlines preprocessing and computational demands. Additionally, the integration of a Gaussian-weighted merging process improves segmentation reliability by reducing block-boundary artifacts, emphasizing the importance of architectural innovations that propagate both local and global contextual information for precise delineation of anatomical structures [57, 85].

In parallel, digital pathology, particularly the registration of whole slide images (WSIs), has witnessed rapid methodological

Table 10: Section Objectives: Medical Imaging, Multimodal, and Cross-Modal Analytics

Objective	Strategy/Approach	Measurable Outcome	Clinical/Operational Implication
Multimodal Data Fusion	Intermediate and late fusion architectures	Diagnostic accuracy, F1-score	Improved disease characterization
Cross-Modal Knowledge Transfer	Transfer learning, modality alignment	Domain adaptation error, ROC-AUC	Resource-efficient model generalization
Conceptual Taxonomy	Unified classification of integration patterns	Framework clarity, reproducibility	Accelerated comparative research
Clinical Deployment	Scalable workflow case studies	Deployment rate, real-world utility	Enhanced healthcare accessibility

evolution due to the challenges posed by gigapixel-scale data and staining heterogeneity. Recent techniques, such as multi-scale ring encoders, enable efficient matching of multi-stained serial WSIs, reducing spatial uncertainties to sub-200 μm levels and significantly accelerating processing compared to manual methods. These innovations facilitate tumor heterogeneity quantification and support the optimization of biomarkers like Ki-67, exemplifying the clinical value of methodological advances in enabling personalized cancer diagnosis and treatment [57].

However, despite these technical achievements, the field faces unresolved challenges that impede reproducibility and scalable benchmarking. Large-scale deployment remains constrained by factors including annotation scarcity, pronounced class imbalance, and heterogeneity across imaging protocols and patient populations. Moreover, current dominant approaches—while powerful—are often data-hungry and can exhibit sensitivity to domain shifts, illustrating their limitations when generalizing to real-world settings or rare pathological presentations. Critics have highlighted the need for more robust, data-efficient learning strategies, and for methods better attuned to dataset heterogeneity.

To mitigate these challenges, ongoing efforts include establishing rigorous multi-center collaborations, introducing well-characterized reference controls, and expanding annotated public datasets, all of which are incrementally improving the generalizability of automated diagnostic systems. Nonetheless, further research is warranted in several directions: (1) development of domain-adaptive and semi-supervised methods that can thrive under scarcity of labeled data; (2) robust procedures for quantifying and addressing inter-center heterogeneity; and (3) the creation and adoption of standardized, transparent evaluation benchmarks to facilitate fair comparison and reproducibility.

In summary, automated segmentation, registration, and diagnostic systems are reshaping clinical workflows, yet significant research gaps persist. Future work should prioritize enhancing data efficiency, model robustness, and generalizability, with explicit focus on bridging the gaps between algorithmic advances and practical, large-scale clinical deployment.

5.2 Imaging Data Fusion and Cross-Modal Analysis

The primary objective of this section is to critically survey modern approaches for imaging data fusion and cross-modal analysis, highlighting both the technical innovations and the distinct operational and ethical challenges inherent in integrating heterogeneous imaging modalities. The intent is to provide readers with a coherent understanding of state-of-the-art fusion architectures, real-world deployment scenarios, associated limitations, and persisting open research problems within this domain.

Imaging data fusion involves the integration of information from multiple imaging sensors or modalities—such as visible spectrum, infrared, hyperspectral, and radar—that differ in spatial, temporal, and spectral characteristics. Cross-modal analysis aims to extract complementary features from these sources, thereby enabling more accurate, robust, and context-aware downstream applications.

A wide spectrum of fusion strategies exists, ranging from early (data-level), intermediate (feature-level), to late (decision-level) fusion. Early fusion typically focuses on raw data alignment challenges, while feature- and decision-level approaches leverage advances in deep learning for representation learning and multimodal reasoning. However, despite technical progress, the selection and optimization of fusion architectures remain contingent on the specific application and operational context, with no consensus on a universally optimal methodology.

Transitioning from technical innovation to operational deployment, practitioners face complex tradeoffs among computational efficiency, information gain, and system robustness. Several real-world deployments underscore these tradeoffs, revealing unforeseen issues such as sensor miscalibration, unexpected environmental effects on data quality, and difficulties in harmonizing temporal resolutions. Furthermore, large-scale fusion systems often expose new vulnerabilities—both technical, like error propagation across modalities, and systemic, such as user bias and trust calibration—requiring multidisciplinary mitigation strategies.

From an ethical and regulatory standpoint, integrating cross-modal imaging data introduces additional layers of complexity. For example, fusing biometric and contextual imagery may trigger privacy concerns regulated differently across jurisdictions, and the potential for covert surveillance exacerbates risks for misuse or unintended consequences. Concrete case studies from healthcare and public safety illustrate the real-world legal ambiguities and failures that have prompted regulatory responses, highlighting the urgent need for domain-specific guidelines.

Despite significant advancements, several unresolved research challenges persist. Open questions include: determining principled criteria for modality selection under resource or privacy constraints; designing explainable fusion architectures that support transparency and auditability; and developing standard benchmarks for evaluating the real-world generalization and robustness of fused models. Debates continue regarding the tradeoffs between end-to-end deep fusion—where representations may be less interpretable—and modular approaches that promote explainability but may sacrifice some performance.

In summary, imaging data fusion and cross-modal analysis represent a rapidly evolving but unsettled area, with ongoing disagreements about best practices, regulatory responses, and the balance between competing technical and ethical priorities. Further

research is needed to establish domain-agnostic design principles, scalable operational workflows, and actionable standards for safe and effective cross-modal imaging system deployment.

5.2.1 Fusion of Imaging, Biosignals, Laboratory, and IoT/Behavioral Data.

The promise of medical AI increasingly rests upon its ability to synthesize multimodal data, a necessity arising from the multifactorial nature of disease and the heterogeneity of inputs in real-world healthcare systems. Fusion approaches incorporate data from diverse domains such as imaging modalities (CT, MRI, PET), biosignals (ECG, EEG), laboratory results, patient-reported outcomes, and IoT or behavioral sensor data streams [33, 35, 56, 83].

Multimodal integration consistently demonstrates improvements in diagnostic and prognostic performance, particularly when radiomics features derived from medical images are combined with complementary clinical, laboratory, and sensor-based information [35, 56, 83]. For example, radiomics enables the transformation of medical images into high-dimensional, quantitative features that, when integrated with clinical or genomic data, facilitate more accurate risk stratification and personalized management, such as predicting recurrence in localized clear cell renal cell carcinoma [56]. Radiomics has also demonstrated the capability to correlate quantitative imaging features with gene expression, tumor phenotype, and microenvironment, yielding validated predictive and prognostic biomarkers important in oncology [35]. Integration of these imaging-derived features with other patient data into classifier models surpasses conventional diagnostics in several applications but underscores the challenges involving data reproducibility, sharing, standardization, and workflow integration. Overcoming these challenges requires standardized protocols, multicenter collaboration, and robust health informatics solutions [35].

At the system level, modern hospital-wide big data platforms are designed for real-time integration of clinical, radiological, laboratory, and administrative data. Such platforms, for example, can continuously ingest data from multiple hospital systems, structurally harmonize heterogeneous datasets, and provide scalable analytics and automated processing to support routine deployment of AI-based clinical decision support [78, 94]. Notably, these systems have enabled substantial operational efficiencies: automated self-service data retrieval has significantly reduced retrieval times and increased the volume of data queries, supporting both research and AI-enabled clinical applications [94]. The integration of IoT-based medical data streams further requires the ability to manage data with varying quality and error characteristics, as illustrated by methods that utilize dynamic classification and anomaly response to preserve service reliability and quality metrics [78].

Fusion methodologies developed for medical AI exhibit broad applicability across sectors. Their adaptation in fields such as industrial prognosis, urban monitoring, and energy management has stimulated the development of modular data architectures and advanced governance frameworks that, in turn, inform best practices for medical multimodal analytics [22, 33, 34, 38, 67, 83, 94, 96]. Cross-sector experiences highlight essential components including standardized data protocols, automated orchestration of heterogeneous streams, dynamic error handling, interoperability and security measures, and the establishment of explainable analytical workflows. Collectively, these underscore the critical importance of modular platforms, robust governance, and adequate computational resources as prerequisites for the successful translation of multimodal analytic techniques from research into clinical practice.

5.2.2 Uncertainty Management, Missing/Correlated Data Handling, and Integrated Diagnostics.

Effective integration of multimodal data streams requires advanced strategies for uncertainty management, addressing missing values, and understanding correlations between modalities—challenges that are especially pronounced in real-world healthcare environments where data is frequently incomplete, heterogeneous, and subject to errors. Recent methodological advancements addressing these issues include:

Application of Dempster-Shafer theory to quantitatively manage evidence conflicts and derive comprehensive global uncertainty measures [25]. This approach enables explicit quantification of uncertainty and conflict among heterogeneous sources, with fusion rules tailored to complex data environments.

Development of custom loss functions and probabilistic data fusion frameworks that directly incorporate the probabilistic structure of measurement errors and missing data [6, 25, 26, 94]. For instance, diffusion-based probabilistic models can reconstruct full-range responses from sparse, noisy, or heterogeneous sensor measurements with high accuracy and flexibility, demonstrating resilience to typical data limitations in clinical and sensing settings [26]. Hospital-scale data platforms now implement metadata management, data standardization, and automated self-service retrieval to support robust large-scale integration of multisource clinical, laboratory, radiology, and management data, while maintaining scalability and consistency [94].

Utilization of data fusion techniques, such as combining molecular embeddings from various single-task models, to enhance prediction in settings defined by data sparsity or weak inter-modality correlation [6]. These approaches can outperform conventional multi-task learning models under conditions of missing or weakly related data sources, thereby improving analytical performance and clinical applicability. Additionally, systematic evaluations of machine learning models for diagnostic tasks, such as early autism spectrum disorder identification, emphasize careful handling of missing data through imputation, normalization, and feature selection, ensuring robust model performance and generalizability in heterogeneous datasets [77].

Within the context of patient monitoring and IoT-assisted healthcare, approaches such as Proportionate Data Analytics (PDA) enable dynamic classification and adaptive responses to anomalous or incomplete data streams by disentangling measurement errors from natural physiological variation. This leads to improved service response ratios and analytic robustness [8]. Integrated diagnostic frameworks—ranging from machine learning-driven risk models to prognostic solutions that combine radiomics and clinical data—demonstrate that explicit uncertainty quantification and mitigation strategies not only bolster analytical trustworthiness but also enhance the practical clinical value of predictive systems.

5.2.3 Diverse Fusion Strategies (Joint, Late, Early) and Applications.

Operationalizing data fusion in medical domains encompasses a spectrum of strategic models, generally classified into early, joint, or late fusion. The principal characteristics and typical advantages of these strategies are summarized in Table 11.

Table 11: Comparison of Fusion Strategies in Multimodal Medical Analytics

Strategy	Integration Stage	Key Advantages	Typical Applications
Early Fusion	Raw data or low-level features combined prior to algorithmic modeling	Enables learning of shared modality representations from the outset; high potential information synergy	Deep neural network training with joint feature embedding
Joint Fusion	Intermediate layers form shared representations during modeling; cross-modal learning possible	Balances cross-modal synergy and robustness; adaptable attention and feature selection	Cross-modal learning (RNNs, Transformers, NetVLAD); video categorization; multimedia retrieval
Late Fusion	Decision outputs of separate models are combined post-modeling	Enhances robustness to missing or noisy modalities; modular adaptation	Ensemble learning; multi-system decision aggregation

In healthcare contexts, the decision between early, late, and hybrid fusion depends on data characteristics, the extent and nature of missingness or misalignment, and computational constraints. For example, cross-modal learning frameworks that utilize unbiased retrieval risk estimators and complementary contrastive learning approaches demonstrate strong performance even with noisy or partially mismatched modality pairings, such as those arising in image-text datasets. Techniques that emphasize negative sample regularization are particularly effective at reducing overfitting to noisy, incorrectly labeled positive pairs, while harnessing all available negative pairs to mitigate underfitting in weakly supervised contexts [36, 47, 51]. Furthermore, self-paced multi-view co-training strategies, which incorporate distributed optimization and allow for scalable, PAC-learnable integration across multiple heterogeneous views, provide an extension of classical semi-supervised learning to address the complexity and scale of contemporary medical and IoT scenarios. These frameworks offer robust mechanisms for data selection and co-regularization, improving both generalization and learning efficiency [27, 42, 53, 63].

Across these diverse applications, multimodal data fusion is propelling significant advances in patient monitoring, precision diagnostics, and rapid medical multimedia search—catalyzing a broader transition toward holistic, data-driven clinical decision support [8, 28, 30, 31, 33, 34, 36, 47, 61, 63, 77, 87]. Despite ongoing challenges related to generalizability, interpretability, and standardization, continuous methodological innovations are charting a path toward more robust, trustworthy, and fully integrated medical AI. The confluence of advanced segmentation, multimodal data integration, and uncertainty-aware analytics is ultimately shaping the future landscape of precision medicine.

6 Operational Analytics, Population Health, and Clinical Deployment

This section aims to analyze the real-world implementation of multimodal AI systems in healthcare, focusing specifically on operational analytics, the impact on population health, and challenges associated with clinical deployment. Building on the technical innovations discussed in preceding sections, we shift focus here to the translation of these advances into practical, scalable, and ethically responsible healthcare solutions. The section further clarifies open research questions, unresolved challenges, and competing considerations in regulatory, ethical, and technical domains.

Operational analytics encompasses methods for harnessing multimodal data to monitor, evaluate, and optimize healthcare workflows and outcomes. The section examines the modalities and architectures deployed in current practice, contrasting their integration strategies and practical trade-offs. Critical comparative analysis highlights how fusion approaches affect downstream analytics and decision-making, emphasizing consensus gaps and areas of methodological divergence. The narrative then extends to population health applications, identifying the unique challenges in synthesizing structured and unstructured datasets across diverse patient cohorts. Discussions highlight both the advances and the persistent limitations in achieving robust, generalizable insights at scale.

Clinical deployment introduces an additional layer of regulatory, ethical, and technical complexity. Real-world deployments often expose unanticipated fragilities, including data privacy risks, cross-population biases, and issues of interpretability and trustworthiness for end users. Notable regulatory frameworks, such as HIPAA in the US and GDPR in the EU, provide concrete legal boundaries but may conflict with the scalability of large-scale data fusion or synthetic data generation. The section examines exemplar deployment failures and unresolved debates—such as the balance between accuracy and explainability or the ongoing lack of consensus in synthetic data validation standards—that continue to shape health AI adoption.

Despite significant progress, open problems remain in building robust, reproducible, and defensible multimodal systems for health operations and population management. These include: reconciling competing stakeholder priorities when designing analytics pipelines; configuring federated or privacy-preserving fusion architectures to satisfy regulatory demands; and developing systematic evaluation frameworks for synthetic data reliability in clinical settings. In particular, there is no established consensus on optimal cross-modal fusion strategies or on metric selection for validating real-world impact, with differing methodologies yielding tradeoffs in scalability, interpretability, and generalizability.

To summarize, this section articulates key operational opportunities and deployment risks in multimodal health analytics. It underscores the need for further research into unified evaluation standards, comparative benchmarking across fusion methods, and more transparent, ethically robust system architectures. Persistent open challenges and areas of methodological debate highlight the necessity for both technical innovation and interdisciplinary dialogue as the field transitions toward real-world clinical adoption.

6.1 AI-Driven Hospital and Population Health Operations

Objectives: This subsection aims to delineate how AI and IoT are reshaping hospital operations and population health management, with a focus on improving resource utilization, enabling real-time data-driven interventions, advancing data integration platforms, and addressing persistent equity and quality challenges.

The convergence of artificial intelligence (AI) and Internet of Things (IoT) technologies is fundamentally transforming operational paradigms in hospitals and population health management, enabling data-driven resource optimization, real-time surveillance, and tailored interventions. In hospital environments, conventional surgery scheduling approaches often result in suboptimal resource utilization and workflow inefficiencies due to insufficient consideration of patient-specific and procedure-specific attributes. With

the advent of machine learning, particularly multivariate ensemble models, there has been a marked shift: these systems now harness pre-procedural clinical and administrative data to transcend the limitations of historical mean-based and manually corrected methods. The outcome is a significant reduction in root mean squared error for time predictions and a notable decrease in late-running cases for both elective and acute surgeries. Such advances have immediate clinical ramifications—more precise scheduling directly leads to reduced surgery cancellations, improved operating room throughput, and enhanced predictability in perioperative care processes, collectively contributing to the alleviation of longstanding bottlenecks that negatively impact clinical and economic outcomes [72].

Beyond the operating suite, institution-wide big data platforms are emerging as foundational to the integration of diverse data sources—including clinical, laboratory, and administrative streams—within secure and scalable infrastructures. The West China Hospital Big Data Platform (WCH-BDP) serves as an exemplary model: its architecture is designed for automated, real-time data ingestion and standardized retrieval, harmonizing in excess of 8,000 discrete clinical variables. This robust computational backbone simultaneously advances both operational analytics and AI-driven clinical applications. Critically, the transition from manual to automated data access (reducing retrieval from hours to minutes) underscores the profound potential for such systems to impact not only research productivity but also immediate clinical decision-making, such as critical event detection and real-time management support. These capabilities illustrate the operational benefits of cross-domain, real-time analytics [94].

At the population health level, AI and IoT technologies forge innovative models for epidemiological surveillance and adaptive interventions, yet also highlight enduring challenges. For example, dynamic transmission modeling—combined with detailed economic analysis—demonstrates that focused intervention strategies, such as vector control in high-risk zones, are essential for achieving disease elimination goals in a cost-effective manner. However, these models also demonstrate operational fragility as disease prevalence declines, necessitating ongoing adjustments to surveillance and resource allocation in response to fluctuating program coverage and operational limitations. This underscores the imperative for real-time, adaptive data pipelines that support sustainable epidemic monitoring and the judicious distribution of limited resources, particularly in settings with high variability and resource constraints [47, 52]. Comparative approaches, such as biosynthetic multiplexed controls and in silico or CRISPR-edited reference materials, are debated for benchmarking diagnostic assay performance, each offering varying engineerability, cost, and applicability [47]. Emerging synthetic data benchmarking standards represent an active area for future development, especially in evaluating AI diagnostic reliability under real-world constraints.

IoT-enabled ubiquitous health monitoring illustrates both the potential and the complexity introduced by population-scale analytics. The recently proposed Proportionate Data Analytics (PDA) framework enhances the reliability of anomaly detection within heterogeneous health data streams by dynamically categorizing errors and variations. This approach improves the specificity and responsiveness of digital health services in real-world, error-prone environments by continuously adapting to stream quality and user

context. Consequently, PDA represents a critical advancement in robust, quality-aware analytics for scalable remote health monitoring and telemedicine systems [99]. Despite these advances, substantial data-related challenges persist:

heterogeneous noise within multi-source data streams
incomplete or imbalanced modality representation
sample-wise variations in data quality

These factors collectively undermine predictive reliability, especially in settings characterized by multimodal and inconsistent data sources [100]. Alternative multimodal fusion frameworks are increasingly contrasted by their ability to manage such low-quality and variable data, with current debates focusing on methods for robust integration, missing modality handling, and real-time adaptation to heterogeneous environments. Effective mitigation of these challenges is essential for the equitable and practical deployment of such analytics at scale.

The interface between digital health platforms and AI analytics is exemplified by large-scale, application-driven interventions in chronic disease management. Deployments integrating continuous glucose monitoring (derived from wearables), dietary and physical activity tracking, and personalized feedback mechanisms have yielded measurable improvements in glycemic control, weight reduction, and behavioral engagement among diabetic and prediabetic cohorts. Notably, these interventions operate autonomously—without dependence on synchronous human coaching—demonstrating the feasibility of scalable population outreach at minimal marginal cost and offering a strategic approach to longstanding gaps in diabetes prevention. However, extending these models across heterogeneous populations necessitates careful consideration of cultural, demographic, and behavioral diversity, alongside stringent long-term validation and thoughtful integration with existing healthcare pathways [5].

Despite rapid advances in digital health capabilities, substantial disparities in care delivery remain entrenched. Analyses of large-scale diagnostic service utilization reveal that, even as technological improvements narrow the access gap, pronounced inequities persist in the timeliness and thoroughness of follow-up care across racial and socioeconomic strata. This observation highlights the necessity of embedding technological innovation within a comprehensive digital equity framework, which addresses not only technical availability but also the accessibility and cultural appropriateness of health interventions [52, 70].

The digitization and interoperability of health information systems have progressed significantly due to targeted national policies and infrastructure investments. Currently, a majority of U.S. hospitals and prescribers possess the capacity for electronic data exchange and public health reporting. Nonetheless, persistent barriers—including technical, usability, and equity-related constraints—particularly impact smaller and rural providers, limiting the fluidity of data necessary for patient-centered, population-level analytics. Overcoming these divides requires continuous enhancement of data standards, development of transparent and trustworthy AI models, and an unwavering commitment to user-centric design principles that facilitate actionable insights at the clinical point of care [10, 94].

Key Takeaways:

Main technical and operational goals include resource optimization, real-time surveillance, adaptive data pipelines, and equitable deployment of AI-driven analytics.

AI-powered scheduling and big data platforms (e.g., WCH-BDP) demonstrably improve operational throughput, data integration, and analytics for both clinical and managerial tasks.

The choice of benchmarking strategies and frameworks (biosynthetic controls, synthetic/in silico data, robust fusion models) is the subject of ongoing technical debate, with significant implications for reliability and privacy.

Persistent data quality heterogeneity, incomplete modalities, and inconsistent multi-source inputs remain active barriers requiring specialized mitigation strategies.

Equity gaps persist despite widespread digital adoption, necessitating ongoing focus on both access and culturally competent care.

Continued progress depends on transparent AI, improved standards, and infrastructure that bridges urban–rural and socioeconomic data divides.

6.2 Clinical Decision Support and Human-in-the-Loop Analytics

Section Objective: This subsection examines the technical and operational goals underpinning modern clinical decision support (CDS) and human-in-the-loop analytics, with a focus on integrating explainability, usability, and scalable analytics platforms into clinical workflows to enhance safety, effectiveness, clinician trust, and equitable care delivery.

The steady advancement of operational analytics and population health management is intricately connected to progress in CDS systems, where user-centric design and analytic transparency are critical determinants of success. Irrespective of technical sophistication, AI solutions must be intentionally developed for integration with clinician workflows to achieve their potential in delivering safe, effective, and equitable care. Usability is pivotal in this context: systems that impose cognitive overload, lack transparency in reasoning, or disrupt established workflows are likely to encounter diminished adoption or even active resistance, notwithstanding their superior algorithmic performance [39, 77].

To address these barriers, explainable AI (XAI) frameworks have become indispensable in the pursuit of trustworthy CDS. Methods such as saliency mapping, Shapley Additive Explanations (SHAP), and Local Interpretable Model-Agnostic Explanations (LIME), once confined to academic research, are now operational in clinical contexts. These tools deliver granular, case-specific illumination of the diagnostic logic underpinning complex deep learning and ensemble models, thereby empowering clinicians to understand, scrutinize, and ultimately trust AI-driven decisions. Notable examples include time-series analysis for arrhythmia detection using XAI for quasi-periodic biomedical signals [39] and disease classification in diabetes with a self-explainable interface leveraging SHAP for model transparency [20], where explanatory tools have rendered model outputs accessible and actionable for both clinicians and patients [8].

The integration of analytics with IoT systems further depends on accommodating heterogeneity in data sources and user preferences. In the context of AI-assisted neuroimaging, workflows must support both the delivery of precise quantitative biomarkers and facilitate expert oversight and correction, given the notable costs of annotation and the presence of systematic biases in large-scale training datasets. While multimodal fusion remains on the frontier of clinical analytics, the challenges of pervasive data noise, missingness, and intricate interdependencies among modalities demand the development of highly adaptive, transparent, and robust architectures to ensure clinical validity and broad generalizability [8, 100].

At the organizational scale, analytic usability is substantively improved by platforms that balance democratized access to data and models with stringent governance and security. These platforms enable clinicians and operational leaders to derive and act upon real-time insights, supporting the evolution toward self-service, data-driven workflows. One exemplary platform, the WCH-BDP, achieves integration of large-scale, heterogeneous medical data while ensuring robust security and governance, leading to notable efficiency and application benefits for clinical and operational leaders [94]. At the infrastructure level, proportionate data analytics has been used to manage heterogeneous medical data streams in IoT-based healthcare systems, differentiating data by variation and error to optimize clinical responses and reliability [78]. Such workflows underpin the human-in-the-loop analytics paradigm, wherein continuous expert feedback calibrates, audits, and refines AI models for local applicability, mitigating risks of model drift and emergent bias [78, 94].

Ultimately, successful CDS deployment requires meticulous alignment with end-user requirements and an explicit focus on clinician trust. Sustainable implementation is attainable only through dynamic, iterative feedback loops between AI developers and clinical stakeholders, ensuring that usability, actionability, and analytic transparency are all practically realized at the point of care. This symbiotic relationship between humans and AI is not simply a matter of user preference; it constitutes a foundational precondition for the delivery of reliable, safe, and ethical analytics in patient care and population health initiatives alike [20, 39, 78].

Key Takeaways for Clinical Decision Support and Human-in-the-Loop Analytics:
- Effective CDS requires seamless integration with clinical workflows, prioritizing usability and analytic transparency across all stages of deployment [77, 94].
- Explainable AI (XAI) frameworks such as SHAP, LIME, and saliency mapping are increasingly operationalized, delivering actionable insights and supporting clinician trust [8, 20, 39].
- Multimodal data fusion in clinical analytics demands adaptive, robust architectures to overcome pervasive noise, missing data, and intermodal dependencies, especially in real-world settings [8, 100].
- Advanced organizational platforms (e.g., WCH-BDP, PDA) balance large-scale data access, strong governance, and self-service analytics, paving the way for scalable, resilient, and secure deployments [78, 94].
- Human-in-the-loop paradigms, emphasizing continuous expert input and calibration, are essential to ensuring long-term clinical applicability, safety, and fairness of AI-driven models [78, 94].

7 Synthetic Data Generation, Privacy, and Security

This section critically analyzes the technical and operational goals, methodologies, and multifaceted challenges associated with synthetic data generation, with particular attention to privacy, security, benchmarking, and ethical implications. By explicitly outlining section objectives, we aim to help readers quickly identify technical and regulatory priorities as well as major open debates.

The primary technical objectives in this domain are: enabling data augmentation when original data is scarce or restricted, safeguarding privacy and mitigating bias, and ensuring that synthetic data meets fidelity requirements for downstream applications without introducing unacceptable privacy or security vulnerabilities. Operational goals encompass supporting efficient and scalable deployment, while also adhering to evolving regulatory standards and building public trust in data use.

Synthetic data generation has become a critical component of modern data-driven AI systems. It offers concrete pathways for data augmentation, privacy preservation, and bias mitigation. Its operational significance is pronounced when real data is limited, sensitive, or subject to legal restrictions. Integration of synthetic data introduces new dimensions to model training, validation, and deployment pipelines, directly influencing issues of data provenance, utility, and fairness.

Transitioning from core technical advances to deployment, synthesizing data not only affects downstream model accuracy and generalizability but also raises nuanced questions regarding consent, data ownership, and exposure to novel security threats. In sensitive sectors such as healthcare or finance, complexity increases due to the heightened need for robust privacy guarantees alongside utility.

From a legal and ethical perspective, regulations like the General Data Protection Regulation (GDPR) define strict boundaries for processing and sharing personal data, demanding clear criteria for anonymity and re-identification risk. Failure to sufficiently anonymize or assess privacy threats risks regulatory violations and erodes public trust. There is broad disagreement over definitions of de-identification and "sufficient" privacy, resulting in variant methodologies and ongoing debate within the field.

The landscape of synthetic data privacy metrics is particularly contested. Some frameworks prioritize formal privacy guarantees (e.g., differential privacy), while others weigh empirical tests for re-identification risk. Benchmarking standards are still emerging, with initiatives aiming to establish transparent and reproducible comparisons across data generators, privacy metrics, and fidelity benchmarks. Comparative discussion of these alternatives highlights the absence of consensus and the need for rigorous evaluation criteria.

Concrete failures and case studies—such as unintended leakage from de-identified health records and synthetic data enabling adversarial attacks—demonstrate that even well-intentioned protocols may yield unforeseen vulnerabilities or discriminatory outcomes, especially under adversarial pressure. These failures underscore the technical and ethical precariousness inherent in current practices.

Persistent technical limitations include the challenge of balancing data utility with enforceable privacy guarantees, addressing

robustness and explainability of data generators, and minimizing risks such as model inversion or memorization. Deep generative model approaches in particular demand ongoing trade-offs: boosting fidelity can inadvertently reveal identifying information, while maximizing privacy protections may diminish utility for downstream tasks. No single standard exists for quantifying privacy leakage or benchmarking fidelity, as different approaches often produce divergent results.

Operational security hazards—such as data poisoning, inversion attacks, and adversarial manipulations—necessitate constant vigilance as deployment environments become more complex. Integrating privacy-preserving mechanisms, notably differential privacy, into generation pipelines remains a prominent open research direction, complicated by practical implementation and verification hurdles.

In sum, synthetic data generation occupies a critical intersection of technical innovation and ethical responsibility, with unresolved challenges around regulatory compliance, technical guarantees, and managing trade-offs between privacy, utility, and operational security. Empirical evidence and high-profile case studies argue for the development of more rigorous guidelines, frameworks, and transparent reporting standards, underlining the importance of cross-disciplinary collaboration in mitigating systemic risks.

Key takeaways for this section are as follows:

Technical and operational goals: balancing fidelity, privacy, and scalable deployment. Major debates: absence of consensus on synthetic data benchmarking standards and privacy risk metrics. Current limitations: trade-offs between data utility and privacy; lack of explainability, robust metrics, and practical privacy-preserving techniques. Operational challenges: vulnerability to adversarial attack, implementation obstacles for privacy mechanisms, and evolving regulatory requirements. Open directions: need for transparent comparative frameworks, cross-disciplinary solutions, and more comprehensive evaluation methods.

7.1 Generation Approaches and Clinical/IoT Use Cases

The application of synthetic data within healthcare and IoT-driven environments is propelled by the imperative for privacy-preserving analytics, augmentation of data-scarce scenarios, and the facilitation of regulatory compliance for both clinical and real-time health monitoring contexts [29, 78]. The inherent difficulty in accessing large-scale, real-world patient data—due to legal, proprietary, and privacy constraints—poses significant barriers to effective algorithm development, benchmarking, and deployment across diverse patient populations. Synthetic data provides a pragmatic solution by generating realistic but non-identifiable datasets that enable open exchange, expedite the iterative cycle of hypothesis testing, and support bias analysis. Furthermore, such data can serve as controlled testbeds for operational procedures and regulatory stress-testing [29, 78].

Technologically, the generation of synthetic data for healthcare and IoT applications employs a range of advanced techniques meticulously tailored to replicate the complexity and diversity of medical information. Among the most prevalent frameworks are Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs),

and state-of-the-art natural language processing (NLP) models, which underpin the synthesis of electronic health records (EHRs), medical imaging datasets, and heterogeneous multivariate IoT sensor streams [29]. Agent-based modeling further extends capabilities by emulating intricate temporal patient interactions and disease trajectories—a crucial element for simulating clinical trials or conducting epidemiological studies. The continuous evolution of these methods now encompasses the generation of time-series data, unstructured clinical text, genomic profiles, and cross-modal datasets, aligning with the emerging requirements of precision health monitoring and digital twin paradigms [29, 78].

Key application domains for synthetic data include privacy-preserving, multicenter AI model training, augmentation of data for rare disease cohort analyses, simulation of IoT environments for robust edge-device validation, and policy development and planning at governmental and institutional levels.

Collectively, synthetic data is gaining centrality in both centralized and federated analytic workflows—enabling privacy, scalability, and inclusivity across the rapidly evolving digital health ecosystem [29, 78].

7.2 Challenges Ethics, Legalities, and Technical Hurdles

Despite the clear benefits, the adoption of synthetic and real-time data streams within healthcare and IoT domains is hindered by a complex intersection of ethical, legal, and technical challenges that directly impact utility, trust, and societal acceptance. One of the most prominent concerns is the perpetuation or amplification of data biases: generative models trained on incomplete, under-representative, or demographically skewed datasets may yield synthetic records that reinforce clinical inaccuracies or institutional inequities, thus compromising algorithmic fairness and posing risks to marginalized populations [29, 52]. Moreover, the opacity of many generative processes complicates auditability, making it difficult to ascertain data provenance, validate representativeness, or perform in-depth post-hoc analyses—functions that are vital for regulatory oversight, transparency, and quality assurance [53, 61, 78, 94].

Privacy risks remain salient even with supposedly de-identified synthetic data. Recent advances in re-identification techniques—including membership inference and adversarial linkage attacks—have raised credible concerns that synthetic datasets that appear realistic may, in fact, unintentionally expose sensitive attributes or facilitate indirect identification of individuals [29, 78, 79]. While legal frameworks such as HIPAA (USA) and GDPR (Europe) set out foundational requirements for privacy, they frequently lack clear, harmonized mechanisms for synthetic data, especially in the context of real-time or streaming IoT scenarios. Consequently, organizations must contend with a patchwork of technical safeguards and jurisdiction-specific compliance requirements, navigating a landscape complicated by distributed data sources, disparate standards, and evolving regulatory guidance [6, 8, 10, 11, 22, 24, 26–30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 79, 89, 90, 92–95, 99, 100, 102].

Differential privacy (DP) methods—such as Private Aggregation of Teacher Ensembles GAN (PATE-GAN)—offer strong mathematical privacy guarantees against information leakage; however, their adoption is limited by mathematical complexity, potential reductions in data utility, and the significant computational resources required [8, 29]. Implementing DP in resource-constrained, heterogeneous IoT environments introduces further challenges related to the need for real-time compliance monitoring and adaptive threat responses [6, 8, 10, 11, 22, 24, 26–30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 79, 89, 90, 92–95, 99, 100, 102].

The regulatory landscape further compounds these difficulties. Disparities in definitions, enforcement, and regulatory scope across national and supranational levels create obstacles for harmonized risk management, while most established frameworks remain tailored to static, offline data rather than to continuously generated or streaming IoT data. This gap exposes real-time systems to threats such as sensor spoofing, data poisoning, and unauthorized data linkages—vulnerabilities inadequately addressed by current guidelines [6, 10, 11, 24, 26–30, 35, 41, 46, 47, 51–53, 58, 60, 61, 64, 75, 92–95, 98, 102]. The absence of harmonized definitions, standardized data interoperability, and robust audit mechanisms ultimately undermines trust among data controllers, practitioners, regulators, and the general public [10, 24, 29, 94].

A structured comparison of major technical and governance challenges is provided in Table 12, highlighting the interplay and current mitigation limitations.

7.3 Advanced Governance and Secure Infrastructures

This subsection aims to provide a unified perspective on the evolving frameworks and persistent challenges in establishing robust governance and security infrastructures for synthetic, clinical, and IoT-derived health datasets. Our survey's overarching goal is to synthesize recent technical, regulatory, and ethical advances, identifying how these domains interact to ensure data quality, privacy, accountability, and trustworthy deployment in complex digital health environments.

Overcoming these intertwined challenges demands systemic transformation in governance, security, and auditability of health data. Contemporary frameworks emphasize end-to-end "digital chain-of-custody" infrastructures, integrating cryptographic tools for provenance tracking, digital signatures, and audit logs, alongside machine-readable access control policies. These mechanisms promote accountability spanning initial data generation to analytical sharing, enforce regulatory compliance, and enable traceability [1, 6, 7, 11, 12, 14–16, 21, 26–29, 35, 41, 43, 46, 47, 51, 60, 61, 65, 66, 75, 78, 79, 84, 91–93, 95, 98, 102]. In support of transparency and traceability, Giuffrè et al. [29] explicitly call for a chain-of-custody model to enhance patient trust and regulatory oversight across the data lifecycle in synthetic health data.

Technical innovations such as blockchain and distributed ledger technologies further support governance by introducing a tamper-evident layer to record-keeping, distributing control among stakeholders, and enabling programmable, granular access. These features are especially valuable in federated and cross-institutional environments, where boundaries of trust are ambiguous, thus securing scalable analysis across interconnected IoT networks [1, 6, 7, 11, 12,

Table 12: Overview of Major Challenges in Synthetic Data for Healthcare and IoT

Challenge Category	Manifestation in Synthetic Data Contexts	Current Limitations of Mitigation Strategies
Bias and Fairness	Propagation or amplification of demographic and clinical inaccuracies; persistent care disparities [52]	Limited model transparency; insufficient bias auditing tools; difficulty representing rare subpopulations and evaluating impact on marginalized groups
Auditability and Provenance	Opaque generative models hinder source-to-sample tracking and interpretability; challenges in quality control for high-dimensional, heterogeneous data [53, 61, 78, 94]	Absence of standardized audit frameworks; model complexity restricts explainability; lack of traceability in multimodal data streams
Privacy and Re-identification	Susceptibility to membership inference and linkage attacks; confidentiality risks in realistic synthetic samples even after de-identification [29, 78, 79]	Incomplete or patchwork legal coverage; computational and utility tradeoffs with privacy-preserving technologies; technical challenges in federated/streaming settings
Legal and Regulatory Gaps	Fragmented definitions across jurisdictions, static-data-centric frameworks overlook dynamic, real-time IoT realities; lack of interoperability and standardization [10, 24, 94]	Disjointed regulatory oversight; inconsistent enforcement; limited guidance for streaming and cross-border data use

14–16, 21, 26–29, 35, 41, 43, 46, 47, 51, 60, 61, 65, 66, 75, 78, 79, 84, 91–93, 95, 98, 102]. Nevertheless, this decentralization introduces trade-offs, including increased system complexity, consensus latency, and the necessary balancing of transparency (for verifiability and trust) against confidentiality (to safeguard sensitive health information).

Bridging technical and regulatory responses, privacy-preserving analytics such as homomorphic encryption, secure multi-party computation, and privacy-aware machine learning enable computation on encrypted or obfuscated data [1, 6, 7, 11, 12, 14, 15, 21, 26–29, 43, 46, 47, 51, 60, 66, 78, 84, 91, 95, 98]. In IoT settings, edge intelligence and decentralized storage confine risk to device-local or micro-environmental domains. As indicated in Roth et al. [79], such advances expand the attack surface and complicate endpoint management, making continuous re-evaluation of compliance with legal regimes (e.g., GDPR, HIPAA) necessary—particularly as device and network boundaries blur.

To synthesize the technical and governance dimensions and to provide a formal conceptual model, we propose an integrative taxonomy (see Table 13) that classifies key approaches and challenges across three axes: *Data Lifecycle Domain*, *Enabling Technologies*, and *Governance Objectives*. This taxonomy serves as a bridge between disparate solutions and highlights areas needing further harmonization.

Short transition statements that bridge technical and regulatory discussion: Moving from technical solutions to their regulatory implications, the adoption of such tools must be coupled with ongoing adaptation of legal and ethical standards, frequently outpaced by the rapid development of new modalities and data flows. For example, federated learning methods [79] enable privacy-preserving collaborative AI in healthcare but also introduce complex compliance and trust challenges across institutional and national boundaries.

A critical, ongoing challenge is thus the dynamic reconciliation of technical innovation and evolving regulation. Governance structures need to persistently balance data utility and patient privacy, optimize open data sharing for model quality while reinforcing access controls for legal and ethical compliance. Regulators are tasked with keeping pace with technical advances, resolving cross-jurisdictional inconsistencies, and specifying clear standards for algorithmic transparency, auditability, and fairness [29, 78].

Transparency and equity must be embedded throughout these frameworks. Best practices urge interdisciplinary collaboration–uniting technologists, clinicians, policymakers, ethicists, and patient advocates–to co-create governance models, monitor emerging risks, and iteratively refine controls. Such collaboration should also anticipate new risks to data and model quality: adversarial attacks on chain-of-custody, provenance spoofing, and drift from privacy transformations are recognized concerns. Methodological advances are needed in areas such as automated verification for provenance, robust consensus protocols, and context-aware privacy controls balancing individual risk [29, 79].

In summary, advanced governance and secure infrastructures constitute an evolving interplay between technical, regulatory, and ethical strategies. Their sustained development and alignment are essential for ensuring trust, accountability, and high-quality deployment in digital medicine. For improved reader traceability, a mapping of grouped citation keys to core/seminal advances is provided in Appendix A.

8 Data Quality, Benchmarking, and Technical Robustness

Section Objectives: This section aims to critically examine how data quality, benchmarking, and technical robustness intersect to underpin trustworthy, accountable, and governance-aligned AI systems. We focus on analyzing the technical challenges and trade-offs involved, how these dimensions support or constrain regulatory compliance and privacy goals, and what methodological innovations are needed to bridge gaps between technical and governance requirements. Throughout, we endeavor to clarify transitions among these subtopics and highlight the need for an integrated conceptual framework.

Ensuring data quality is foundational for the reliability of AI systems. High-quality data not only reduces bias and errors but also supports meaningful benchmarking and compliance with regulatory standards. However, maintaining rigorous data standards often conflicts with the need for large-scale data collection, leading to trade-offs between dataset size, representativeness, and annotation accuracy. In many deployments, the lack of transparent data provenance and insufficient documentation exacerbates these issues, making it difficult to assess long-term robustness or to compare models fairly. This tension is particularly pronounced in contexts with strict privacy and governance requirements, where the need for data minimization and secure handling can further constrain data utility.

Transition: From data quality, we move to the challenges surrounding benchmarking practices, which serve as the locus where technical measurements, regulatory demands, and privacy requirements converge.

Benchmarking practices similarly face critical limitations. While standardized benchmarks facilitate comparative evaluation, they may inadvertently encourage overfitting to particular datasets and fail to capture evolving real-world complexities. Moreover, disparities in benchmark adoption can obscure the generalizability of experimental results. Integrating robust governance and privacy mechanisms within benchmarking protocols remains an unresolved challenge; comprehensive benchmarks must account not only for performance metrics but also for system security, fairness, and data protection measures.

Transition: Given the dual role of benchmarking as both an evaluation tool and a governance checkpoint, we now examine how technical robustness must evolve in light of these requirements.

Table 13: Taxonomy Bridging Technical and Governance Strategies in Digital Health Data Infrastructures

Data Lifecycle Domain	Enabling Technologies/Processes	Governance Objectives/Challenges
Acquisition & Generation	Consent management, edge devices, secure sensors	Data minimization, provenance, access legitimacy
Storage & Transmission	Cryptographic protocols, distributed ledgers, federated learning	Data integrity, jurisdictional compliance, scalability, auditability
Analysis & Sharing	Homomorphic encryption, MPC, privacy-aware ML, differential privacy	Utility vs. privacy, transparency, fairness, bias mitigation
Deployment & Monitoring	Audit logs, automated compliance checks, continuous monitoring	Adaptive regulation, traceability, risk oversight, quality drift

Technical robustness involves resilience against distributional shifts, adversarial attacks, and implementation faults. To address these issues, current methods employ techniques such as adversarial training, uncertainty quantification, and rigorous stress testing. However, these approaches can increase system complexity, introduce computational overhead, and may trade off against other desiderata such as interpretability or ease of deployment. Effective technical robustness demands holistic integration with governance and privacy frameworks, yet many existing systems treat these domains in isolation. More explicit cross-domain strategies are needed to ensure that privacy-preserving or governance-driven constraints do not inadvertently introduce new vulnerabilities or degrade system performance.

Transition: Acknowledging the divides outlined above, we explore open challenges that suggest the need for unified frameworks to align technical and governance dimensions.

Open challenges in this domain motivate the need for innovative methodological advancements. Chief among these is the development of adaptive, context-aware quality control and benchmarking procedures that can flexibly balance technical, privacy, and governance objectives. For instance, dynamic benchmarks capable of evolving alongside deployed systems could help mitigate the risks of benchmark overfitting and foster more realistic assessments of robustness. Likewise, methodologies for verifiable data quality auditing, privacy-preserving metric computation, and automated governance compliance checking present opportunities for substantive progress.

Toward an Integrated Taxonomy: To address the compartmentalization identified above, we propose considering a unified taxonomy for data quality, benchmarking, and technical robustness that classifies solutions by: (1) their technical scope (e.g., data, model, system); (2) their governance interplay (e.g., compliance, auditability, privacy guarantees); and (3) the nature of assessment (static vs. dynamic, holistic vs. modular). Such a taxonomy can guide both research and deployment toward harmonized, governance-aware technical practices.

In summary, the interplay among data quality, benchmarking, and technical robustness reveals numerous open challenges—ranging from reconciling data utility with privacy constraints to aligning governance requirements with robust, scalable evaluation methods. Addressing these issues requires not only methodological innovations but also deeper integration across typically compartmentalized areas, paving the way for more trustworthy and accountable AI systems.

8.1 Data Quality in Clinical and IoT/Streaming Contexts

8.1.1 Problems of Data Heterogeneity, Class Imbalance, and Missing Modalities in Biomedical and IoT Data. Emerging applications of artificial intelligence in healthcare, biomedicine, and the Internet of Things (IoT) increasingly depend on heterogeneous, multimodal datasets. Biomedical data encompass structured clinical records, high-dimensional medical images, physiological signals, and continuous streams from wearable or IoT sensors. Each data source displays distinct characteristics, such as varying sampling frequencies, noise patterns, and degrees of completeness. This variability complicates effective integration and generalization, directly challenging robust clinical translation and impeding the practical deployment of AI models in real-world settings [10, 24, 100]. As recent surveys highlight, platforms designed for large-scale integration of multisource healthcare data must address structurization, standardization, and privacy while supporting high-throughput analytics across diverse modalities [94]. Models built on uniform or single-source data typically fail to generalize within diverse and often noisy clinical environments, underscoring the need for rigorous data integration and governance [24, 94].

Class imbalance remains a persistent barrier in both clinical and IoT-driven data. In practice, adverse or critical health events (for instance, seizures or acute disease onset) represent only a minority of samples compared to a predominance of routine or normal signals. This imbalance can bias learning algorithms to favor majority classes, lowering sensitivity for detecting rare but clinically important events [24, 94]. In addition, real-world datasets frequently suffer from missing values or entire missing modalities, attributed to device malfunctions, inconsistent patient engagement, or disruptions in data collection. IoT-based healthcare amplifies these concerns, with data streams exposed to intermittent loss, asynchronous arrival, corruption, and heterogeneous device standards. These present substantial obstacles to efficient preprocessing, integration, and inference, demanding both innovative technical strategies and improved platform architectures [8, 78].

Such quality issues are often interdependent; their intersection gives rise to complex failure modes. Multimodal fusion is particularly susceptible to missing modalities: absence or corruption of even a single data source can degrade downstream performance unless compensated by robust imputation or modality-aware methods [100]. Comprehensive surveys and large hospital deployment experiences consistently stress these interacting data factors as central barriers to successful deployment of machine learning pipelines in biomedical and IoT domains [78, 94, 100]. Overcoming these challenges will necessitate not only rigorous technological progress in

data integration and standardization but also transparent reporting, governance, API enhancements, and equitable access models to realize the full potential of AI-powered healthcare solutions [24, 94].

8.1.2 Technical Preprocessing: SMOTE, Artifact Rejection, Denoising, and Anomaly Detection for IoT and Wearable Streams.

Multiple sophisticated preprocessing pipelines have been developed to address the persistent challenges found in biomedical and streaming data, notably concerning class imbalance, noise, and artifact contamination. The Synthetic Minority Oversampling Technique (SMOTE) and related algorithms are widely adopted for addressing class imbalance in tabular and stream data. By generating synthetic samples of underrepresented classes, SMOTE facilitates improved detection of rare but clinically essential events [37, 77]. These benefits, however, are constrained in highly nonstationary streams or ultra-high dimensional applications, where SMOTE may lead to overfitting or the creation of unrealistic sample boundaries if not carefully applied [78].

Artifact rejection and denoising constitute foundational steps in preprocessing physiological signals, especially in real-world, uncontrolled recording environments. Classic filtering approaches, thresholding, imputation, and dimensionality reduction using principal component analysis remain common. Recent advances have seen the application of deep learning-based approaches such as convolutional autoencoders and dual-stream neural networks, which offer robust capacity to discriminate context-dependent artifact signatures and complex noise profiles from physiological signals [37, 71, 80]. For example, multiresolution CNN frameworks that extract features from decomposed frequency bands [37], and dual-polynomial decomposition for optimized signal denoising [80], have demonstrated both improved artifact mitigation and preservation of operationally or clinically important signal content. Novel decompositions (e.g., Hurst method for real-time EEG analysis [71]) further enable extraction of relevant features under dynamic recording conditions.

In IoT and continuous streaming environments, real-time anomaly detection techniques such as Proportionate Data Analytics (PDA) are increasingly used to dynamically segment incoming data, discriminating anomalies caused by device errors or data corruption from genuine physiological changes, and enabling actionable responses in health monitoring systems [8, 78]. PDA leverages dynamic error modeling and stream variation analysis through regression-based segmentation, supporting responsive service delivery in healthcare IoT architectures.

Despite these technical advances, a persistent challenge remains in balancing the selectivity of denoising with the risk of discarding clinically informative subtle signals, or conversely, leaving residual confounders that impair downstream analysis. State-of-the-art pipelines increasingly incorporate artifact-aware loss functions, explicit uncertainty modeling, and adaptive preprocessing tailored to both signal characteristics and target clinical endpoints [78]. This integration is crucial for high-stakes biomedical contexts where both false negatives (signal loss) and false positives (artifact retention) are consequential.

As summarized in Table 14, each preprocessing approach exhibits context-dependent advantages and limitations, emphasizing the ongoing need for tailored solutions that integrate uncertainty

modeling and adaptive methods to address the complexities of biomedical and IoT data streams.

8.2 Benchmarking and Validation

8.2.1 Multicenter, Real-World, and Robust Benchmarking Approaches.

Rigorous benchmarking is fundamental to the development and validation of generalizable machine learning models in biomedicine and IoT. While single-center, retrospective studies have historically been prevalent, the field is shifting toward multicenter and real-world validation paradigms. This evolution is motivated by recurring failures of models that perform well under limited, homogeneous conditions but falter upon deployment in diverse clinical environments [1, 41, 92]. Initiatives such as the Image Biomarker Standardization Initiative (IBSI) exemplify the critical importance of harmonizing feature definitions and preprocessing protocols, thereby enabling reproducibility, transparent reporting, and robust validation across different software platforms [53, 55, 61].

Beyond imaging, the creation of large-scale, multicenter data platforms that encompass multimodal clinical and streaming information has become instrumental for assessing algorithmic generalizability [19, 24, 28, 75, 84, 88, 102]. These endeavors foster new benchmark datasets that incorporate real-world complexities—including missing data, fluctuating class distributions, and heterogeneous device characteristics—facilitating stress-testing of algorithms before clinical or operational deployment [24]. Contemporary benchmarking protocols now emphasize open-source curation, transparent metrics, and prospective, temporally-anchored validation, ensuring that reported performance is not merely an artifact of retrospective overfitting [1, 41, 53].

8.2.2 Addressing Computational/Deployment Limitations, Dataset Imbalance, and Patient Variability.

In practice, benchmarking exposes a convergence of technical and systemic obstacles that must be overcome to enable scalable AI deployment. Computational limitations persist, as deep learning models—despite their strong performance on curated datasets—often require resources exceeding those available in typical clinical or edge computing environments. Real-time analysis of high-resolution streaming data is thus frequently infeasible without model optimization. To address these constraints, research increasingly focuses on neural network quantization and the development of computationally efficient architectures, which are essential for practical edge deployment in both clinical and IoT settings [2, 18, 37, 78].

In addition, data imbalance and patient variability introduce further challenges. Achieving truly robust benchmarking requires evaluations that are stratified across demographic, clinical, and device-specific subgroups. This includes explicit quantification of performance disparities and the application of non-inferiority analyses for underrepresented populations [77, 78]. Addressing these challenges involves algorithmic innovations such as domain adaptation, meta-learning, and sample reweighting, infrastructure improvements for federated, multi-institutional data aggregation, and the adoption of statistical methodologies for subgroup-specific performance assessment [18, 37]. For example, seizure prediction and detection studies have systematically tackled dataset imbalance with approaches like cluster-based under-sampling and SMOTE, enabling high sensitivity while containing false positive rates [18, 37].

Table 14: Comparison of Preprocessing Techniques for Biomedical and IoT Data

Technique	Primary Target	Key Advantages	Limitations
SMOTE	Class imbalance	Enables rare event detection; robust for tabular or balanced low-dimensional data	Limited for high-dimensional, nonstationary streams; risk of unrealistic instances or overfitting
Filtering/Thresholding	Structured noise, artifacts	Computationally efficient, easy to interpret and deploy	May discard relevant features; less adaptive to complex or context-specific artifact sources
Deep Learning Denoising (Autoencoders, dual-stream nets, multiresolution CNN, MRDPI)	Complex, non-stationary noise/artifacts	Learns representative features; adaptable to data complexity; improved artifact/signal separation [37, 71, 80]	High data and computational requirements; challenges in interpretability and generalization
Anomaly Detection (PDA)	Device faults, data corruption in streams	Supports real-time error modeling and dynamic segmentation; maintains service reliability in IoT healthcare [8, 78]	Vulnerable to context-ambiguous anomalies; performance dependent on underlying error model

Table 15 summarizes reported results from recent representative works, highlighting variations in sensitivity and false-alarm metrics for state-of-the-art models operating on distinct EEG datasets.

8.2.3 Ensuring Translation to Scalable and Interoperable Platforms. Even when technical performance metrics are met, the broader translational impact of AI systems in healthcare and IoT depends fundamentally on integration within scalable and interoperable platforms. The proliferation of heterogeneous health IT infrastructures has led to fragmented data ecosystems, impeding efficient data exchange, aggregation, and systematic benchmarking [28, 35, 92, 102]. Recent progress—including advances in interoperability standards, terminology harmonization, and development of federated system architectures—has begun to address these silos, supporting more seamless deployment of AI solutions across distributed settings [35, 97, 102]. Persisting barriers, however, comprise non-standardized data formats, privacy and compliance constraints, heterogeneous computational capacity, and limited adoption of open APIs. Overcoming these hurdles is essential for realizing robust, scalable, and equitable AI-enabled healthcare and IoT ecosystems [28, 75].

9 Key Challenges, Open Problems, and Future Directions

This section synthesizes the primary challenges, articulates open problems, and delineates key opportunities for methodological advancement across the field. The objectives here are: (1) to provide an integrated analysis of outstanding technical, privacy, and governance challenges; (2) to expose and categorize current research gaps; and (3) to foster new perspectives by introducing a conceptual taxonomy reflecting cross-domain interdependencies.

We identify three major axes along which the landscape of challenges and future work can be analyzed: Technical Robustness and Scalability, Privacy and Ethical Governance, and Practical Integration & Deployment. These axes characterize the core domains of obstacles and serve as the basis for the following cross-domain taxonomy:

A salient research challenge lies in balancing these often competing axes, as advances in one dimension (e.g., technical scalability) can exacerbate difficulties in another (e.g., privacy risks). By organizing the discussion according to this taxonomy, we facilitate the identification of critical tradeoffs and illuminate emerging research gaps spanning multiple domains.

Looking forward, a key direction involves deepening the methodological integration across technical, ethical, and practical challenges. Future research may benefit from adopting holistic frameworks that explicitly incorporate multi-stakeholder perspectives and support transparent cross-domain analytics, particularly in fields such as healthcare and IoT, where data, regulatory, and operational constraints intersect most acutely.

This taxonomy is intended to guide both researchers and practitioners in navigating the multifaceted landscape by clarifying priorities, supporting comparative assessment, and encouraging cross-disciplinary innovation.

9.1 Section Objectives

The objectives of this section are threefold: (1) to critically analyze the main technical, regulatory, and governance-related obstacles faced in the field; (2) to contrast diverse approaches and their respective tradeoffs; (3) to map out how these open challenges can shape and inspire future methodological innovation.

9.2 Critical Analysis and Contrasting Perspectives

One of the central technical challenges involves balancing model performance with privacy and governance constraints. For example, advanced machine learning methods often require access to large volumes of high-quality data, yet privacy-preserving techniques (such as differential privacy or federated learning) may reduce data fidelity or complicate data integration. This tension requires deliberate architectural choices, where increasing privacy or robustness can negatively impact accuracy or efficiency. Furthermore, there is a persistent gap in regulatory alignment, as governance frameworks (such as GDPR or industry-specific guidelines) may lag behind technological progress. These frameworks impose additional requirements that can either impede or shape technical development. Thus, careful assessment of tradeoffs between openness, transparency, and compliance is integral to system design.

A further perspective pertains to deployment: integrating privacy and governance solutions directly into data pipelines and model life-cycles is nontrivial. Often, technical solutions are retrofitted instead of being considered from the outset, which can introduce inefficiencies or lead to suboptimal outcomes. In addition, rapid innovation cycles in machine learning can outpace regulatory responses, which may create uncertainty and hinder deployment in sensitive domains.

9.3 Open Problems and Their Impact on Methodological Advancement

Several open problems are evident and point to promising research avenues. First, there is a need for frameworks that reconcile privacy preservation with minimal performance degradation, particularly in distributed or federated environments. Second, more work is needed on adaptive governance mechanisms that can dynamically incorporate evolving regulatory requirements into technical workflows.

Data quality and representativeness continue to be critical: advances in mitigating dataset bias, augmenting underrepresented modalities, and validating data provenance are essential for robust

Table 15: Performance benchmarking of recent seizure prediction and detection models on representative EEG datasets.

Dataset	Sensitivity (%)	False Positive Rate	Reference
CHB-MIT	82	0.058	[37]
Kaggle	85	0.19	[37]
State-of-art (CHB-MIT, 13 pts)	90.2	0.0713	[37]

Table 16: Conceptual Taxonomy of Key Challenges and Open Problems

Axis	Description	Representative Challenges	Future Directions
Technical Robustness and Scalability	Ensuring accuracy, efficiency, and resilience in real-world, large-scale deployments.	Model generalization, adaptation to heterogeneous data, scalability bottlenecks.	Modular architectures; adaptive algorithms; cross-modal learning.
Privacy and Ethical Governance	Preserving confidentiality, fairness, and compliance within evolving policy frameworks.	Secure data sharing, algorithmic bias, transparency, regulatory uncertainty.	Federated learning; privacy-preserving analytics; explainable AI mechanisms.
Practical Integration & Deployment	Bridging developmental innovations with operational use in complex environments.	System interoperability, socio-technical barriers, real-time constraints.	Standardized interfaces; participatory design; edge-to-cloud orchestration.

system performance. Such approaches demand methodological innovation in data sampling, augmentation, and active learning.

Finally, improving the integration of privacy, governance, and technical mechanisms can enable end-to-end trustworthy pipelines. For example, modular solutions that allow dynamic adjustment of privacy budgets or compliance checks during model deployment are underexplored. Addressing these gaps will likely call for advances in interdisciplinary methodologies, as well as tighter collaboration between regulatory stakeholders and technical practitioners.

In summary, the challenges outlined here—balancing privacy, governance, technical scalability, and data quality—are not only barriers but also catalysts for the next generation of methodological contributions. Clearer articulation of these issues, as above, should help motivate and guide ongoing research.

9.4 Technical, Methodological, and Practical Challenges

The deployment of advanced analytics and artificial intelligence (AI) in healthcare, especially within environments utilizing Internet of Things (IoT), real-time monitoring, and resource-constrained settings, is impeded by a range of technical and practical barriers. Foremost among these is data heterogeneity: the coexistence of varied data sources, disparate modalities, and inconsistent formats and data qualities, covering everything from clinical EHRs to biosensor feeds to imaging studies [2–4, 6, 8, 10, 11, 18, 22, 24, 25, 28, 33–38, 41–44, 47–49, 53, 55, 56, 59, 61, 63, 64, 67, 69, 75–79, 83, 84, 86, 88, 91–94, 96, 100, 101]. This diversity complicates the integration process, often resulting in datasets with missing values, inconsistencies, or significant noise, ultimately eroding the robustness and generalizability of downstream algorithms [2–4, 6, 8, 10, 18, 22, 24, 25, 28, 33–38, 41, 42, 44, 47–49, 53, 55, 56, 59, 61, 63, 64, 67, 69, 75–78, 83, 84, 86, 88, 92–94, 96, 100, 101]. Handling these challenges often requires advanced fusion, harmonization, and robust preprocessing pipelines [34, 44, 75, 100].

A parallel obstacle is the limited availability of high-quality annotations. Many machine learning and deep learning models depend on large, well-annotated datasets, which remain rare for certain diseases, underrepresented populations, or novel data modalities such as wearable sensor streams or next-generation molecular diagnostics [28, 53, 56, 69, 96]. Particularly in medical imaging applications, annotation relies on highly skilled human experts through laborious manual labeling processes, creating significant bottlenecks not only for model training but also for reliable external validation. Systematic reviews underscore that small, imbalanced, or single-institution datasets exacerbate these challenges and hinder model generalization to broader clinical practice [53, 61]. For rare diseases or low-prevalence conditions, data scarcity further compounds these issues, impeding model development and evaluation [4, 42, 47, 48, 56]. Device- and protocol-specific variability may further limit the applicability of learned models to new clinical settings [48, 53, 83, 96].

These technical hurdles become even more pronounced as analytic systems are scaled for deployment in real-time or edge environments, including rural or small-clinic contexts where computational infrastructure, memory, and network bandwidth are often highly constrained [3, 6, 8, 22, 24, 25, 33–35, 38, 47, 53, 76, 78, 88, 91, 100]. Modern AI models—especially large, multimodal, or federated architectures—frequently exceed the capacities of typical IoT hardware, which may have limited power sources, minimal storage, and unreliable connectivity. Additional barriers include data synchronization challenges, intermittent streaming, packet loss, high latency, and device failures, all of which demand robust anomaly detection, effective imputation of missing data, and adaptive model retraining methodologies [37, 47, 53, 59, 79, 100].

Recent literature highlights emerging solutions such as unsupervised and semi-supervised learning to mitigate annotation requirements [53, 63, 76, 92], development of edge-optimized or hardware-aware model architectures [18, 37, 53, 100], and the use of federated learning or domain adaptation approaches to improve data privacy and generalizability across settings [41, 63, 79]. Standardization protocols for data curation and interoperability are also being advanced in both imaging and molecular domains to improve consistency and downstream utility [10, 24, 42, 84, 94]. Nevertheless, widespread adoption and reproducibility of these advancements remain limited, and practical deployment of AI in real-world, heterogeneous healthcare environments is still stymied by persistent gaps in scalability, interoperability, and evidence of clinical effectiveness. This underscores a continued need for rigorous methodological research, expanded public datasets, and stronger national and international standards for system integration, annotation quality, and model evaluation.

9.5 Interpretability and Clinical Impact

The effective translation of analytic innovations into clinical practice is contingent upon interpretability, clinical utility, and generalizability. High-performing, yet opaque, models—particularly deep neural networks and large foundation models operating on images, signals, and clinical text—often function as "black boxes." This opacity undermines clinician trust and impedes both validation and regulatory acceptance [2, 8, 20, 22, 25, 28, 32, 34, 35, 39, 53, 54, 56, 61, 62, 64, 67, 78, 82, 100]. Clinical utility is intimately bound to interpretability: beyond providing predictions (such as disease classification or risk scores), models must articulate the underlying rationale, delineating salient image areas or critical features in time-series data. Recent advances in explainable AI (XAI)—including post-hoc saliency mapping, attention visualization, and Shapley value-based feature attribution—enable both global and case-specific interpretative insights, facilitating scientific validation and regulatory review [2, 22, 39, 53, 54, 64, 82, 100].

A distinct but interrelated concern lies in ensuring analytic generalizability across diverse clinical settings, populations, and data acquisition systems. Many models are prone to overfitting to specific institutions, cohorts, or medical devices, resulting in decreased performance in cross-site or prospective validations [2, 25, 28, 44, 56, 62, 78]. This phenomenon is observed across domains, from radiomics to real-time vital sign prediction. Efforts to enhance generalizability include domain adaptation, transfer learning, multi-site federated learning, and incorporation of biological or clinical priors into model regularization. Nevertheless, rigorous external validation and systematic evaluation at scale remain insufficient [2, 25, 34, 39, 44, 54, 78, 100].

9.6 Standardization, Equity, and Ethical Considerations

This integrative section synthesizes cross-domain challenges and opportunities at the intersection of standardization, equity, and ethics in analytic healthcare, reinforcing this survey's broader aim of mapping technical advances, persistent barriers, and future needs across imaging, -omics, and EHR modalities.

Advances in analytic healthcare depend upon multifaceted standardization and an unwavering commitment to equity and ethical principles. Across modalities—including imaging, -omics, and electronic health record (EHR) data—variability in workflows, acquisition protocols, annotation definitions, metadata schemas, and feature nomenclature hinders reproducibility and impedes model transferability [1, 3, 6, 8, 22, 27, 28, 33, 35, 38, 42, 46, 47, 51–54, 56, 61, 69, 78, 83, 84, 94, 96, 99, 100]. International initiatives, such as the Image Biomarker Standardization Initiative (IBSI), have advanced harmonization in specific domains like radiomics [1]. However, non-standardized practices in feature extraction, segmentation, and reporting continue to undermine multi-center reproducibility and clinical translatability [3, 8, 28, 35, 38, 53, 69, 78, 84, 100].

Comparative synthesis across domains reveals parallel struggles in standardizing data practices. In molecular and genomic testing, consensus guidelines stress the need for robust validation and continuous quality control across laboratories and platforms [42, 47]. Big data health platforms for large hospitals also depend on integrated metadata management, terminology harmonization, and longitudinal, cross-system data governance [94]. Machine learning-based predictive analytics and multi-modal fusion face persistent challenges from inconsistent data schemas, sparse and heterogeneous clinical datasets, and lack of agreed benchmarks for evaluation and reporting [8, 53, 78, 100]. These concerns echo in imaging, where protocols for acquisition, segmentation, and biased annotation propagate to downstream clinical or AI deployments [1, 35, 56, 61, 83, 84, 96].

Equity in analytic healthcare extends beyond algorithmic fairness to encompass broader social determinants and the digital divide. Persistent disparities are shaped by infrastructural deficits, inconsistent connectivity, device interoperability limitations, and the geographic maldistribution of expertise and support [5, 8, 10, 28, 47, 52, 70, 78, 94, 99, 100]. Access to diagnostic services, for instance, remains uneven even in technologically advanced healthcare environments [10, 52]. Digital health interventions show promise for chronic disease management, yet may exacerbate or mitigate disparities depending on their real-world accessibility and ongoing support for diverse populations [8, 10, 78, 94, 99]. These issues are compounded by technical complexity, resource limitations, and the lack of tailored governance and transparent algorithms across digital ecosystems [10, 94].

To promote both reproducibility and fairness, several best practices emerge: Clinical workflows and analytic protocols should be harmonized across domains, data types, and geographic locations. Bias mitigation, transparent reporting, and inclusive dataset curation—particularly representing under-represented or high-need groups—are essential. Investments in equitable digital infrastructure, interoperability, and participatory design processes are needed, with special attention to marginalized or underserved populations. Cross-disciplinary initiatives should focus on standard operating procedures, validation practices, and scalable quality assurance frameworks. Collaborative efforts should address persistent barriers to data sharing, privacy, explainability, and inclusive innovation at policy and organizational levels.

These steps aim to mitigate model bias and advance equitable patient outcomes across demographic and resource strata [1, 3, 8, 28, 38, 42, 47, 51, 54, 61, 69, 78, 83, 84, 94, 96, 99, 100]. In the context of increasing reliance on digital health solutions for chronic disease management and pandemic response, these concerns become even more critical.

Despite significant progress, unresolved debates continue to shape the field. For example, the relative merits of federated data harmonization versus centralized standardization remain unsettled, as do the best approaches for balancing privacy and data utility [8, 78, 94]. The literature reflects contrasting strategies and unresolved issues around the biological interpretability of radiomics features [1, 84], multi-institutional reproducibility in machine learning [53, 61, 69], and how best to assure equitable technology diffusion across resource settings [10, 28, 52, 99].

Key Future Research Gaps and Needs: - Development of cross-platform, cross-modality standards for metadata and reporting that can be implemented at scale and adapted to evolving technologies. - Consensus on best practices for multi-center, multi-modal datasets, including robust validation, bias mitigation, and transparent quality assurance. - Interdisciplinary frameworks for measuring, reporting,

and mitigating algorithmic and systemic biases that extend into operational and population levels. - Strategies for addressing the digital divide, including infrastructure investments, participatory design, and governance models tailored to marginalized communities. - Systematic approaches to integrating ethical, regulatory, and social considerations into each stage of analytic health technology life cycles.

Continued cross-domain dialogue, evidence synthesis, and policy engagement are essential to ensuring that advances in data-driven healthcare translate into robust, reproducible, and equitable outcomes.

9.7 Privacy, Security, and Compliance

As analytic and data-sharing platforms become increasingly sophisticated, ensuring privacy, security, and regulatory compliance grows ever more imperative. Legislation such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) mandates stringent controls over data sharing, necessitating advanced technical solutions for privacy-preserving analytics and secure pipeline design [6, 8, 10, 11, 22, 24, 26–30, 33–35, 38, 40, 41, 44–47, 51–53, 58, 60, 61, 64, 73, 75, 78, 79, 89, 90, 92–95, 98–100, 102]. There is growing cross-disciplinary recognition that privacy and security are not merely technical or legal matters, but demand integrated consideration of data lifecycle governance, transparency, and ethical imperatives across domains such as medical imaging [35, 53, 61, 79], health informatics [10, 24, 94], industrial/IoT settings [22, 33, 78], and energy management [33].

Emerging threats encompass data breaches, the risks of re-identification, and the use of synthetic data for privacy enhancement—all of which introduce new attack surfaces and ethical dilemmas, especially when balancing auditability and privacy across diverse data types and stakeholder environments [24, 29, 94]. Techniques such as federated learning and differential privacy are central to technical mitigation; however, they introduce computational burdens and auditability challenges that are yet to be fully resolved [8, 22, 26, 28, 29, 33, 38, 45, 46, 53, 78, 89, 94, 99, 100]. While synthetic data and privacy-preserving computation offer hope for broader data accessibility and reduced regulatory risk, practical implementation is still hampered by challenges in bias, interpretability, and the difficulty of auditing data transformations for compliance [29].

Compounding these considerations is the dynamic and continuous nature of data originating from wearables, IoT devices, and streaming EHR feeds. This necessitates adaptable legal, technical, and ethical frameworks to ensure auditability and accountability as both data and clinical applications evolve under shifting interoperability and stakeholder requirements [6, 8, 10, 22, 28, 29, 34, 38, 47, 52, 59, 78, 79, 94, 100]. Among the foremost challenges are:

Establishing scalable governance for data access and secondary use

Designing fair and comprehensible consent models suitable for ongoing data collection

Conducting regular audits, especially in multi-center and cross-border collaborations

Integrating privacy-by-design and transparency-enabling toolkits from the outset

Despite growing recognition, consistent protocols for compliance verification and accountability remain underdeveloped.

Contrasting Perspectives and Open Debates: Sectors such as health care, industrial prognosis, and building energy management have each evolved divergent approaches to balancing privacy and data utility [22, 24, 33, 79, 94]. In medical imaging, multi-institutional collaboration and federated learning are vital yet challenged by data heterogeneity and legacy privacy concerns [53, 79]. Big data platforms advocate for extensive data integration and streamlined access but note that standardization, security, and equitable patient access remain persistent bottlenecks [10, 94]. By contrast, industrial and energy domains highlight the promise of data fusion, secure analytics, and adaptive consent models, but frequently grapple with questions of transparency, stakeholder trust, and deployment cost [6, 22, 26, 33]. There is debate about whether technical mitigation can ever wholly substitute for strong organizational and legal oversight, particularly regarding explainability, auditability, and risk mitigation in high-stakes settings [24, 29].

Explicit Research Gaps and Future Directions:

Robust methodologies for verifying privacy and compliance in multi-modal, cross-border data sharing pipelines

Comprehensive frameworks for scalable, dynamic consent and secondary data use

Auditable implementations of synthetic data and federated learning in clinical and industrial practice

Practical standards for privacy-by-design and explainability in data pipeline architectures across diverse domains

Unified protocols to support regular, automated audits adaptable to evolving regulatory environments

9.8 Algorithmic Innovation, Architecture, and Recommendations

Section Objectives and Scope: This section explicitly surveys recent advances and open challenges in scalable, ethical, and interoperable analytic architectures for healthcare. We synthesize algorithmic and architectural strategies underpinning modern healthcare analytics—spanning continual and federated learning, hybrid modeling, and self-supervised approaches—with attention to deployment trade-offs and governance. Typical audiences include clinicians seeking translational impacts, AI researchers advancing methods for multimodal/heterogeneous data, and engineers developing clinical platforms. By structuring a unifying taxonomy of analytic design tensions and proposing foundational principles, we also position this survey as a reference for cross-domain practitioners and policy stakeholders.

Innovations in algorithmic design and system architecture are critical to enabling scalable, adaptable, and trustworthy healthcare analytics within the growing complexity and volume of clinical data sources. Driving this progress, recent research emphasizes continual and federated learning, advanced hybrid models, and self-supervised learning to address the intrinsic heterogeneity, high annotation costs, and pervasive decentralization of data across IoT-equipped healthcare ecosystems and intermittently connected environments [3, 6, 8, 22, 26, 28, 33–36, 38, 47, 49, 53, 58, 61, 63, 64,

87, 88, 99, 100]. These algorithmic paradigms increasingly serve as cornerstones for personalized medicine, robust population- and disease-level surveillance, and automatable clinical decision support at scale.

Deployment in practice necessitates reconciling dynamic regulatory, clinical, and technical requirements. More modular and interoperable architectures—prioritizing plug-and-play extensions, data governance, and ethical transparency—are strongly promoted to accelerate safe clinical translation and to align with shifting regulatory mandates [1, 3, 6, 7, 11, 12, 14–16, 21, 22, 25–28, 30, 33–35, 38, 42–44, 47, 49, 51, 52, 54, 55, 58, 60–62, 65, 66, 74, 75, 78, 82, 84, 89, 91–94, 99, 100, 102]. The central tensions—trade-offs among explainability, resource efficiency, interoperability, and ethical governance—are summarized in Table 17 as a foundational taxonomy for analytic architecture design.

A pronounced gap in the literature is the lack of validated frameworks, taxonomies, and real-world benchmarks supporting truly modular, interoperable analytic platforms that guarantee transparency, ethical accountability, and governance across heterogeneous clinical domains. Efforts on data standardization—such as multiplexed NGS reference materials [47], imaging harmonization initiatives [1, 84], and open API-driven integration coupled with robust governance—have increased, yet broader adoption and alignment across sectors remain essential [1, 3, 6, 22, 25, 26, 28, 33–35, 38, 47, 49, 51, 52, 54, 55, 60, 61, 78, 84, 89, 93, 94, 100].

Architectural Blueprint and Unifying Taxonomy: Building on consensus and critical syntheses [1, 8, 28, 34, 47, 61, 84, 88, 94, 100], we propose a unifying architectural blueprint for future healthcare analytics. This blueprint highlights the necessity for:

(1) *Comprehensive Modularity:* Universal modular frameworks that enable plug-and-play adaptation across clinical, genomic, imaging, and sensor data domains [93, 94]. (2) *Standardized Benchmarking:* Real-world, cross-institutional benchmarks and open datasets covering multimodal and low-quality data scenarios [1, 84, 94, 100]. (3) *Transparent Explainability:* Layered explainability from architectural design to output-level interpretation, using recent XAI techniques tailored for regulatory and clinical acceptance [34, 82, 94]. (4) *Context-aware Scalability:* Scalable, resource-adaptable deployments from cloud to edge with mechanisms for continuous learning, data drift management, and local adaptation [49, 58, 61, 64, 78, 99]. (5) *Ethical and Regulatory Compliance:* Integrated, auditable governance (e.g., logging, access control, data lineage) that accommodates varying legal and societal requirements [28, 34, 94]. (6) *Multimodal/Heterogeneous Data Fusion:* Mechanisms for robust, context-aware multimodal fusion coping with incomplete, imbalanced, and noisy data [6, 36, 63, 78, 88, 100].

While notable consensus exists on these priorities, unresolved debates persist regarding the trade-offs between model transparency and predictive accuracy, centralized versus federated deployments, and decisions between data minimization and multimodal fusion for privacy and performance. The literature also highlights context-dependence in the efficacy of analytic strategies; for instance, hybrid models employing continual and federated learning alongside robust self-supervision and cross-modal fusion are often most adaptable to real-world variation [8, 36, 49, 58, 63, 64, 78, 88, 94, 99, 100].

Critical Analysis and Future Research Gaps: Current analytic frameworks are often juxtaposed with established but less

flexible deep learning benchmarks lacking sufficient multimodal, low-quality, or imbalanced data coverage [1, 84, 100]. Recent work advocates for evaluation methodologies that test models under simulated noise, missing data, or incomplete modality configurations—mirroring real-world healthcare complexity—and for establishing comparative protocols across institutions and health systems [1, 61, 78, 84, 94, 100].

Furthermore, there is a growing call for research that bridges highly engineered but opaque architectures with more explainable, context-sensitive algorithms (e.g., combining symbolic and neural methodologies, leveraging foundation models calibrated for medical data [34, 58, 82]). This includes investigating approaches enabling system-level auditability and provenance tracing in increasingly automated and data-driven workflows.

Section Summary and Open Gaps: - Establish universally accepted modular frameworks and real-world benchmarks for cross-domain, interoperable analytic architectures that explicitly support clinical, engineering, and ethical user groups. - Develop explainable, audit-ready algorithms that perform robustly with incomplete, noisy, or imbalanced real-world multimodal data, leveraging recent advances in self-supervised and continual learning. - Advance standardization and harmonization protocols for clinical, omics, imaging, and sensor data, ensuring scalable, context-aware integration across both centralized and distributed deployments. - Foster ethical governance ecosystems incorporating dynamic legal, societal, and regulatory considerations, particularly as AI integration widens in international, multicenter settings. - Investigate hybrid model design harnessing federated, continual, and centralized paradigms for improved trade-offs in scalability, resource efficiency, and local adaptation. - Enable open, transparent benchmarking and data sharing, supporting comparative evaluation and reproducibility for diverse healthcare applications at national and global scales.

Potential systematic research strategies to address these gaps include: expanding public cross-institutional datasets with detailed records of data quality/completeness; developing evaluation protocols that prioritize simulation of realistic multimodal data challenges; designing layer-wise explainability and auditing modules as architectural defaults; and fostering cross-sector alliances to unify standards for ethical, technical, and regulatory interoperability [1, 28, 34, 82, 84, 88, 94, 100].

9.9   Summary

In summary, this survey has explored the objectives and scope of healthcare analytics and AI, focusing on critical challenges such as data heterogeneity, annotation scarcity, system scalability, interpretability, equity, standardization, privacy, compliance, and architectural sustainability. Distinct analytic strategies—ranging from traditional statistical modeling to modern machine learning and deep learning approaches—have been evaluated within clinical, operational, and public health domains. Alternative approaches often offer contrasting strengths; for example, deep learning methods may provide superior predictive accuracy but challenge transparency and interpretability, whereas rule-based systems and classical statistical models facilitate more transparent decision-making processes but may be less scalable to large, unstructured datasets. Ongoing

Table 17: Core Algorithmic Tensions Limiting Healthcare Analytics Deployment

Aspect	Desirable Outcomes	Limiting Trade-offs/Challenges
Explainability	Transparent, trustworthy predictions; regulatory compliance	May reduce accuracy and increase computational burden [34, 82]
Interoperability	Seamless integration with legacy and modern systems	Increased complexity and persistent standardization gaps [1, 44, 89, 93, 102]
Resource Efficiency	Feasibility on edge/IoT devices; sustainable operations	Model compression can reduce accuracy; limited scalability under hardware constraints [61, 78, 94]
Ethical Governance	Responsiveness to legal, clinical, and social mandates	Fragmented standards, limited auditability and benchmarking, variance across jurisdictions [28, 34, 94]

debates revolve around the trade-offs between accuracy and interpretability, centralized versus federated data architectures, and the impact of regulatory frameworks on model deployment in practice.

To foster conceptual clarity, we propose a unifying taxonomy for healthcare analytics and AI, structured along three primary axes: (1) Data Source and Modality (e.g., clinical records, imaging, sensor data), (2) Analytical Approach (statistical models, traditional machine learning, deep learning, hybrid systems), and (3) Application Domain (clinical, operational, public health). This schema enables a systematic comparison of methods and helps illuminate research frontiers where integration and innovation are needed.

Moving forward, systematically addressing research gaps will require (a) developing benchmark datasets and standardized evaluation protocols to ensure reproducibility and comparability, (b) designing interpretable models for complex and heterogeneous healthcare data, (c) innovating privacy-preserving learning algorithms to enable robust multi-site collaboration, and (d) fostering interdisciplinary frameworks that embed ethical, regulatory, and sector-specific requirements at all stages of innovation and deployment. Ultimately, integrative approaches that harmonize methodological transparency, technical innovation, and ethical foresight are essential to realizing the transformative and equitable potential of analytics and AI in medicine.

10 Scalability, Toolkits, Standards, and Analytical Ecosystems

The rapid expansion of Artificial Intelligence (AI) applications across diverse domains has necessitated not only improvements in algorithmic performance, but also rigorous inquiry into the accompanying ecosystems supporting development and deployment. This section aims to achieve the following scoped objectives: (1) synthesize advancements and challenges regarding scalability, toolkits, standards, and analytical ecosystems in AI; (2) propose a unifying architectural blueprint to situate these components within a coherent taxonomy; and (3) critically analyze interdependencies and unresolved issues, with explicit guidance for future research.

We propose a taxonomy that distinguishes four foundational pillars: Scalability (system-level architectures, distributed computation), Toolkits (software platforms and experiment frameworks), Standards (protocols and benchmarks for interoperability), and Analytical Ecosystems (integrated environments for empirical research and deployment). This taxonomy is intended to provide a unified lens for understanding how infrastructures co-evolve with advances in AI, highlighting both technical differentiation and cross-cutting concerns.

We first examine how scalability is achieved through distributed architectures and advanced resource management, outlining current paradigms and their respective limitations. This is followed

by a comparison of representative toolkits, emphasizing their roles in reproducible and efficient experimentation and identifying contrasts with established data science frameworks. Next, we analyze emerging standards, underscoring their necessity for interoperability and traceability, and survey the expanding analytical ecosystems that frame contemporary research and industry practice, including causal relationships between ecosystem maturity and research throughput.

At the conclusion of each major subsection, explicit, bullet-pointed research gaps and open questions are highlighted to guide ongoing investigation. Where appropriate, we discuss how these gaps could be systematically addressed through unified benchmarks, modular architectural extensions, or deeper cross-community coordination, distinguishing the present synthesis from existing surveys.

This integrative approach not only maps the present landscape, but also frames a clear research agenda and positions our survey as a uniquely comprehensive and critical contribution to the evolving field of AI infrastructure.

10.1 Scalability

Objectives and Scope: This subsection aims to systematically analyze scalability in the context of large-scale AI, offering a comparative synthesis of technical strategies, unifying architectural patterns, and future challenges. The goal is to move beyond an enumeration of techniques towards an integrated framework that can guide both practitioners and researchers.

The proliferation of large-scale datasets and increasingly complex AI models has elevated scalability as a central concern in both academic and commercial settings. Resource allocation, parallelization, and dataset partitioning have each emerged as critical technical strategies. Distributed frameworks, such as those enabling parallel training over multiple GPUs or compute clusters, are now standard practice, greatly reducing time-to-insight for model development.

To provide a clearer theoretical synthesis, we propose a unifying taxonomy for scalability solutions in AI systems, organized by architectural and operational domains:

Across these categories, scaling challenges intensify when addressing issues such as heterogeneous hardware, variable network topologies, and dynamic workload balancing. Recent research emphasizes the causal interplay between system-level bottlenecks (e.g., network latency, non-uniform memory access) and algorithm design, motivating deeper integration between distributed systems theory and AI workloads.

In comparison to established frameworks, current distributed training paradigms have advanced the state of practice but still demonstrate key limitations with respect to hardware diversity and workload adaptation. There is a nascent shift toward architectural blueprints that integrate self-adaptive scheduling and heterogeneity-awareness, though these remain under-explored at scale.

Table 18: A unifying taxonomy for healthcare analytics and AI approaches

Data Source/Modality	Analytical Approach	Application Domain
Clinical records	Statistical models	Clinical (diagnosis/prognosis)
Imaging	Machine learning	Operational (workflow/patient flow)
Sensor/wearable data	Deep learning	Public Health (epidemiology)
Multimodal/fused data	Hybrid (combined/hierarchical)	Cross-cutting (decision support, policy)

Table 19: Unifying Taxonomy of Scalability Strategies in AI Systems

Category	Description	Typical Approaches	Scalability Challenges
Resource Allocation	Efficient distribution of compute, memory, and storage	Dynamic scheduling, cloud provisioning	Resource contention, multi-tenancy
Parallelization	Executing tasks concurrently across hardware	Data/model parallelism, pipeline parallelism	Synchronization, straggler effects
Dataset Partitioning	Dividing data for distributed processing	Sharding, mini-batch sampling	Data skew, consistency
Distributed Training	Coordinated training across nodes	Parameter servers, all-reduce, federated learning	Communication overhead, fault tolerance
Adaptive Algorithms	Dynamic adjustment based on workload	Online load balancing, elastic scaling	Heterogeneous hardware, real-time requirements

Open Research Gaps and Systematic Approaches: How can current distributed training paradigms be adapted for increasingly heterogeneous hardware environments? Systematic approaches include leveraging hardware abstraction layers and co-designing algorithms with architecture-specific optimizations.

What novel algorithms can address the bottleneck of inter-node communication as model and dataset sizes continue to grow? Promising methods include asynchronous updates, communication compression, and decentralized coordination, each of which present unique trade-offs for accuracy and scalability.

How might dynamic resource scaling be optimized in response to fluctuating workload demands in real-time AI applications? Research directions involve predictive autoscaling, reinforcement learning-based orchestration policies, and closed-loop feedback mechanisms to continuously refine allocation decisions.

This survey uniquely synthesizes these perspectives into an explicit taxonomy and highlights avenues for cross-fertilization between distributed systems and AI research, aiming to catalyze further progress towards robust, scalable AI deployments.

10.2 Toolkits

Objectives and Scope: This section aims to critically examine the evolution, architecture, and challenges of major open-source AI toolkits, with an emphasis on unifying principles, reproducibility strategies, and the emerging need for modularity and portability. We further propose a taxonomy for categorizing toolkits by architecture and intended use, laying groundwork for systematic comparison and highlighting avenues for research.

AI research and deployment are profoundly shaped by the availability of robust, open-source toolkits. These frameworks abstract away much of the underlying complexity associated with hardware, low-level numerical operations, and distributed computing, thus lowering barriers to entry and fostering innovation. Prominent toolkits now offer extensive support for automated differentiation, optimized backends, and seamless integration with cloud platforms.

To facilitate a more systematic understanding, we propose a taxonomy based on three interrelated axes: (i) *Architectural Modularity* (degree of decoupling between core execution engine, user-facing API, and third-party extension points); (ii) *Deployment Portability* (extent of native support for local, distributed, and cloud/hybrid operation); and (iii) *Reproducibility Mechanisms* (built-in support for versioning, environment management, and pipeline encapsulation). Table 20 outlines exemplar toolkit categorizations under this lens.

However, the evolution of toolkits also exposes tensions around version compatibility, long-term support, and the standardization of APIs. Further, reproducibility remains a concern, especially in the context of complex pipeline dependencies and evolving environmental dependencies. Comparative analysis reveals that toolkit architectural choices often trade off extensibility and performance against stability and ease of migration. Additional work is needed to distill causal mechanisms by which toolkit architecture influences experimental replicability and user productivity.

Future research gaps and open questions include: - What best practices and technical mechanisms can further advance reproducibility and robustness in toolkit-based AI experimentation? Approaches could include enhanced isolation of dependencies, automated snapshotting, and unified environment descriptions, building on the taxonomy outlined above. - How might toolkits provide better support for modular, reusable components adaptable across diverse application requirements? Systematic integration of standardized APIs and plug-in architectures remains an open challenge, meriting synthesis from both theoretical and engineering perspectives. - In what ways can toolkits facilitate seamless migration between local and cloud-based development environments? Opportunities exist for toolkits to provide transparent context adaptation and formalized model packaging, further lowering barriers to scalable, reproducible AI research.

By structuring the discussion around this taxonomy and critically contrasting toolkit approaches, this survey seeks to clarify both the landscape and the open pathways for increasing rigor, interoperability, and scalability in AI toolkit design.

10.3 Standards

Objectives: This subsection aims to critically examine the evolving landscape of AI standards, with a focus on key technical, regulatory, and ethical challenges. It synthesizes recent initiatives, identifies trade-offs between standardization goals, and highlights open questions to guide future work.

Table 20: Representative taxonomy of AI toolkits by architectural modularity, deployment portability, and reproducibility mechanisms.

Toolkit	Architectural Modularity	Deployment Portability	Reproducibility Mechanisms
Toolkit A	High (separable core/APIs)	Native support (local/cloud)	Integrated versioning, env encapsulation
Toolkit B	Medium (partial separation)	Limited to local, extensible	Manual version tracking
Toolkit C	Low (monolithic)	Local only	Minimal support

The adoption of standards in AI has become increasingly important for ensuring interoperability, safety, and regulatory compliance. Standards affect not only file and data formats but also model exchange, benchmarking protocols, and documentation requirements. Active initiatives seek to harmonize practices across toolkits and platforms, reinforcing both scientific rigor and commercial viability.

However, the path toward effective standardization is shaped by significant technical and regulatory challenges. Rapid advances in both hardware and software frequently outpace the development of universal standards, leading to fragmentation across competing platforms and diverse application domains. Furthermore, the increasing demand for transparency and reproducibility often intersects with legal and ethical considerations, such as the need to protect intellectual property and data privacy. For example, the debate over common benchmark creation is not only a matter of technical specification, but also of ensuring the inclusiveness and fairness of evaluation across modalities and geographies, with regulatory frameworks placing further constraints on data usage and reporting.

A critical comparison of existing initiatives reveals diverging approaches: some emphasize maximal openness in documentation, while others restrict disclosure to protect proprietary techniques. Regulatory efforts likewise vary between promoting interoperability versus ensuring strict data sovereignty. These methodological differences underscore ongoing controversies about how standards can best accommodate a range of stakeholders, from open-source communities to industry consortia. While certain domains (e.g., medical AI) prioritize rigorous documentation for safety-critical applications, broader sectors may value flexibility even at the cost of cross-platform compatibility.

Transitions across these technical and regulatory fronts illustrate the complexity of establishing robust AI standards. Tackling one challenge—such as harmonizing benchmarks—can introduce further friction with privacy or proprietary interests. This underlines the need for architectural blueprints or taxonomies that systematically reflect the dynamic interplay among benchmarks, privacy, and analytic protocols.

Future research gaps and unresolved questions remain central to advancing the field: - How can standardization practices keep pace with rapidly evolving methods and hardware without stifling innovation or causing obsolescence? - What frameworks are needed to balance the competing demands for transparency, intellectual property protection, and regulatory compliance in standard documentation? - Which new or revised benchmarks can be established to better represent emerging, high-risk AI domains and address

stakeholder concerns? - How might a unifying taxonomy or architectural model inform the next generation of AI standardization efforts across novel application areas?

Summary: The development of robust AI standards is shaped by fast-paced technical progress, regulatory diversity, and complex ethical trade-offs. Systematic approaches to reconciling benchmarks, privacy, and documentation remain a key frontier, underscoring the necessity for more agile, inclusive, and critically informed standardization frameworks.

10.4 Analytical Ecosystems

This section aims to synthesize current knowledge on AI analytical ecosystems, identify foundational challenges, and propose a unified blueprint to support scalable, reproducible, and impactful research. We seek to clarify the organizational landscape, establish a comparative taxonomy, and highlight actionable future directions for advancing open and extensible analytical infrastructures.

AI analytical ecosystems comprise integrated toolchains, software libraries, and collaborative platforms that support the end-to-end lifecycle of research and deployment. These ecosystems enhance productivity, foster collaboration, and facilitate the aggregation and analysis of large-scale experimental results. Increasingly, they integrate visualization tools, experiment tracking systems, and interconnected data management pipelines.

Despite notable progress, significant obstacles remain in unifying disparate ecosystem components, managing data provenance, ensuring reproducibility at scale, and promoting open science across institutional boundaries. Based on our synthesis, we propose the following taxonomy of analytical ecosystem components and their interrelations:

This taxonomy provides a reference structure to guide systematic ecosystem development and evaluation.

Future research gaps and open questions include: How can analytical ecosystems be further streamlined for multi-institutional, cross-disciplinary collaboration? What advances are required in provenance tracking and experiment management for large-scale reproducible research? Which incentives and architectures best promote community-driven growth and sustainability of open analytical ecosystems?

To address these gaps, systematic approaches are needed: the adoption of common data and experiment metadata standards to improve interoperability and provenance; modular, API-driven architectures promoting plug-and-play extensibility across platforms; and community governance frameworks that incentivize contribution, foster transparency, and ensure long-term sustainability. Evaluation of ecosystem effectiveness should include measurable

Table 21: Taxonomy of Core Components in AI Analytical Ecosystems

Component Category	Functionality	Integration Challenges	Typical Solutions
Data Management	Ingestion, storage, versioning, access	Data heterogeneity, provenance tracking	Distributed storage, data lakes, metadata layers
Experiment Tracking	Logging, lineage, results linkage	Reproducibility, scalability	Open-source trackers, automated logging tools
Visualization	Data/metric presentation, interactive exploration	Real-time integration, usability	Dashboard frameworks, visualization libraries
Collaboration	Multi-user management, shared environments	Security, cross-institutional support	Cloud platforms, permission systems
Deployment Pipeline	Continuous integration, model serving	Workflow orchestration, monitoring	CI/CD tools, containerization

objectives such as reproducibility benchmarks, cross-institutional adoption rates, and scalability metrics.

By systematically addressing these infrastructural pillars, guided by the above taxonomy and actionable objectives, the AI community can accelerate the transition toward scalable, reproducible, and trustworthy analytical ecosystems that support both research excellence and efficient deployment across diverse domains.

10.5 Key Datasets and Software Toolkits

The exponential escalation of multi-modal, clinical, and Internet of Things (IoT) data volumes has necessitated a robust foundation of centralized datasets and open-source software toolkits. These developments underpin both methodological advances and translational applications across healthcare and analytical domains. Among the most influential centralized datasets is UniMod1K, which offers synchronized RGB, depth, and language data—addressing the longstanding challenge of multi-modality alignment—thus supporting advanced fusion models for object tracking and monocular depth estimation [85]. ImageNet-ESC, extending the paradigmatic ImageNet dataset into the audio-visual sphere, further promotes few-shot learning by leveraging cross-modal associations and is extensively employed for benchmarking adaptation capabilities in both academic and industrial settings [55, 83]. Each of these datasets reflects domain-specific innovations: UniMod1K targets the integration challenges pertinent to vision and language, while ImageNet-ESC exemplifies the value of cross-modal adaptation in scenarios such as medical image segmentation and endoscopic diagnosis [55, 58, 83].

In digital health, the MIT-BIH Arrhythmia database and CHB-MIT seizure corpus provide vast, annotated electrocardiogram (ECG) and electroencephalogram (EEG) datasets. These resources are invaluable for developing and validating machine learning algorithms, fostering reproducible studies in arrhythmia detection and seizure prediction, respectively [3, 8, 13, 17, 19, 37, 38, 48, 58, 77–79, 86, 96]. Notably, contemporary studies employing MIT-BIH highlight the role of pre-trained convolutional neural networks coupled with sophisticated time-frequency transforms (e.g., continuous wavelet transforms) to maximize classification accuracy for arrhythmic events, emphasizing both performance and clinical interpretability [86]. For example, Song and Lee [86] demonstrated that transforming ECG signals into 2D images via time-frequency methods such as continuous wavelet and short-time Fourier transform enables the use of pre-trained networks like ResNet and DenseNet, achieving up to 96.17% accuracy in arrhythmia classification. Their comparative analysis underscored the effectiveness of continuous wavelet transform (notably the Ricker wavelet) and provided insights into the impact of pre-training and hyperparameter choices when using the MIT-BIH dataset for binary classification. Similarly,

the CHB-MIT and Kaggle's intracranial EEG datasets enable the development of advanced deep learning systems—such as multiresolution convolutional neural networks leveraging frequency band decomposition—for real-time, patient-specific seizure forecasting in clinical and home environments [18, 37, 78]. Ibrahim et al. [37], for instance, detail a seizure prediction approach using multiresolution CNNs and frequency band analysis, reporting sensitivities of 82% (CHB-MIT) and 85% (Kaggle) and demonstrating the importance of patient-specific pipelines and prospective deployment in IoT-based settings. Chatzichristos et al. [18] further highlight the practical success of multimodal, wearable pipeline development for at-home epilepsy management, showing enhanced seizure detection and artifact rejection through multimodal postprocessing with IMU sensor data. Collectively, this ecosystem of curated datasets and evolving evaluation benchmarks facilitates algorithmic benchmarking, enables rigorous cross-institutional comparisons, and underpins the progress of scalable, generalizable AI in healthcare and sensor-driven analytics.

Software toolkits have concomitantly evolved to address the growing heterogeneity and complexity of analytics challenges. The `scikit-multimodallearn` library exemplifies general-purpose, open-source design tailored for multi-modal supervised learning, offering seamless integration with the broader Python machine learning ecosystem and providing specialized formatting algorithms for handling heterogeneous input types [13, 18]. This design supports workflows ranging from classification and regression to exploratory multi-modal analytics, thereby enabling both foundational research and rapid prototyping across disciplines. Radiomics platforms—particularly those stemming from the Image Biomarker Standardization Initiative (IBSI)—play a pivotal role in standardizing the extraction and interpretation of high-dimensional features from medical images across modalities such as CT, PET, and MRI, thereby substantially improving reproducibility and diagnostic reliability [1, 13, 35, 36, 63, 84]. These platforms have catalyzed large-scale studies in oncology, cardiology, and neurology, facilitating harmonized methods for imaging-based phenotype discovery and cross-center benchmarking. Furthermore, open-source code repositories associated with recent publications increasingly supply pre-trained models, annotation utilities, and reproducible computational pipelines that lower barriers to clinical and industrial adoption in areas such as cross-modal analytics, federated medical image analysis, sensor integration, and health monitoring [19, 48, 49, 78, 79]. The contemporary ecosystem is further strengthened by the growing emphasis on modularity, scalability, and high-quality documentation, underpinned by well-specified interfaces and reference datasets, ensuring alignment with the technical demands of large-scale deployment.

Beyond data and software, the rigor of annotation and evaluation remains foundational for robust analytics—particularly in high-stakes domains such as clinical validation and IoT-driven automation. Standardized metrics, including positive percentage agreement (PPA), positive predictive value (PPV), response and delivery times, variation factors, and delivery accuracy, are now routinely utilized to enable fair benchmarking and to monitor performance against established baselines [1, 27, 28, 35, 42, 47, 51, 53, 54, 61, 76, 78, 84, 85, 99]. These metrics are embedded within regulatory frameworks, as demonstrated in next-generation sequencing and imaging analytics, where stringent, error-based workflows involving reference controls and multi-center validation are required for reliability, reproducibility, and auditability [1, 42, 47, 53, 61, 84]. Contrasting perspectives persist regarding key evaluation challenges. Heterogeneity of annotation formats, class imbalance, inconsistencies in ground truth assignment, and variability in domain-appropriate metric selection continue to pose obstacles to model generalizability and clinical translation [27, 28, 47, 51, 78, 99]. For instance, harmonization efforts in radiomics have highlighted the tradeoff between reproducibility and the need for flexible, task-specific feature sets [1, 35, 84]. In the context of EEG and ECG analytics, debates continue regarding the optimal balance between automated and expert-generated labels, and handling domain shifts across institutions and populations [18, 37, 86]. These unresolved aspects underscore an ongoing need for multi-institutional, diverse dataset curation, method transparency, and careful selection of performance metrics tailored to both translational relevance and regulatory standards. Continued attention to these challenges will be critical as the field advances toward automated, interoperable analytics and clinical decision support systems, driving further convergence of domain-specific and general-purpose strategies.

10.6 Secure, Scalable, and Interoperable Analytical Ecosystems

The expansive proliferation of distributed data sources—including federated clinical repositories and diverse IoT sensor streams—has crystallized the demand for analytic ecosystems that are secure, privacy-preserving, interoperable, and efficient at scale. Distributed and federated analytic approaches have emerged as primary solutions to the privacy constraints inherent in health and IoT data. These approaches enable collaborative model training without necessitating cross-institutional raw data exchanges [7, 11, 12, 14–16, 21, 41, 43, 46, 60, 65, 66, 75, 91–93, 95, 98, 102]. For instance, federated learning strategies such as FedAvg and FedProx have demonstrated near-centralized performance in distributed medical imaging tasks, all while maintaining robust privacy guarantees [41, 79]. These frameworks have been evaluated on heterogeneous datasets across multiple modalities and pathology domains, highlighting strong generalizability. However, they also reveal persistent limitations, including inter-site data heterogeneity, communication bottlenecks, and challenges synchronizing model updates [10, 46, 75, 94, 95].

Security and auditability are increasingly enhanced by the integration of blockchain and related ledger technologies, which offer transparent, tamper-evident transaction records of both data access and analytic workflow execution [7, 12, 14, 16, 21, 41, 62, 65, 91, 93].

Combining blockchain-backed data provenance, secure model aggregation, and robust API-level integration enables the construction of verifiable analytical pipelines, which are crucial for regulated clinical environments and IoT deployments alike [1, 6, 26–28, 35, 47, 51, 61, 66, 78, 79, 84, 102]. The merger of federated analytics and blockchain architectures addresses pressing concerns such as data sovereignty, tamper resistance, and multi-institutional coordination. Nevertheless, these benefits often come at the cost of increased system complexity, transaction latency, and heightened resource demands, necessitating careful trade-off analysis in real-world settings [1, 8, 78].

The assurance of efficiency, reproducibility, and regulatory compliance in large-scale, interconnected deployments demands not only technical innovation but also robust standardization and governance frameworks [22, 26–28, 30, 33, 34, 38, 47, 49, 51, 53, 54, 58, 78, 79, 92, 94, 95, 99]. Landmark initiatives—for example, the IBSI project for radiomic feature harmonization and major efforts toward electronic health record interoperability—have yielded essential consensus protocols guiding data formatting, algorithmic benchmarking, and metric reporting, thereby fostering reproducibility and translational potential [1, 13, 30, 36, 84]. Comprehensive data platforms such as WCH-BDP exemplify these principles through their integration of interoperable data lakes, rigorous governance with standardized terminologies, natural language processing-based data structuring, and layered security frameworks, all designed to enable rapid and secure analytics at scale [64, 94]. Notwithstanding these advances, the reality of deployment continues to be shaped by a variety of ongoing challenges:

Incomplete or imperfect standardization across data sources and analytic processes

Persistent data silos and mismatched ontologies hindering integration

Technical heterogeneity across institutions limiting interoperability

Disparities in resource allocation impacting scalability

Evolving requirements for multi-stakeholder governance and consensus

The range of domains where secure and scalable analytical ecosystems are critical extends across healthcare, industrial Internet-of-Things (IoT), scientific research, and energy infrastructure. In healthcare, the focus is often on integrating multimodal clinical, omics, and sensor data under stringent privacy and regulatory regimes. Recent work demonstrates both the promise and complexity of federated learning and big data platforms for clinical data integration and predictive analytics [8, 53, 64, 79, 94]. In industrial and energy domains, data fusion approaches leverage heterogeneous sensor streams and real-time machine learning for prognosis, monitoring, and optimized control, although technical and governance-related hurdles remain a challenge [6, 22, 26, 33, 38]. Across all settings, interoperability issues—arising from mismatched ontologies and technical standards—remain persistent, with significant disparities in infrastructure and resource allocation between institutions or stakeholders [10, 94].

There is also active debate within the literature regarding best practices and long-term strategies for balancing privacy, scalability, and utility. For example, while federated analytics and blockchain integration show promise for privacy and auditability, they present

new trade-offs in computational efficiency and system complexity [1, 8, 78, 79]. Contrasting perspectives concern the extent to which technical harmonization (such as feature definition or algorithm standardization) versus governance (such as data-sharing agreements and oversight) should be prioritized. Furthermore, while standardization is often cited as foundational for reproducibility and translational impact, critics point to the practical limitations of omnibus protocols in fast-evolving technological domains and argue for more adaptive regulatory and technical frameworks [30, 53, 94, 95].

Superimposed on these infrastructure issues is a pressing requirement for analytical transparency, explainability, and trustworthiness, prerequisites for both clinical adoption and regulatory acceptance. Techniques such as comprehensive model auditability, embedded conceptual knowledge representation, graph-based causal modeling, formal verification, and explainability have been identified as foundational for future-proofing analytic pipelines and ensuring compliance within medical AI and IoT ecosystems [28, 34, 49, 53]. Despite their critical importance, these approaches frequently encounter real-world barriers: high annotation overhead, complexity of deployment, and challenges to computational scalability all temper their immediate impact in operational settings. Addressing these limitations requires sustained collaborative efforts spanning technical, clinical, and governance domains, including embedded standards and co-design with end-users. Such efforts remain pivotal for building impactful, sustainable analytics infrastructures across health, scientific, and industrial sectors.

A comparative perspective suggests that while technical solutions are increasingly sophisticated, there is no universally optimal strategy: trade-offs in privacy, efficiency, interoperability, and trustworthiness must be continuously revisited and aligned with domain-specific objectives, available resources, and evolving governance standards. Open debates and unresolved issues persist around the most effective blend of federated, blockchain, and data fusion approaches, particularly as the volume, diversity, and sensitivity of multimodal data continue to grow.

Actionable Takeaways: Practitioners must carefully weigh domain-specific privacy, efficiency, and interoperability requirements, and adopt hybrid strategies as necessary. Technical harmonization must be coupled with robust governance to ensure scalability and trust. Continual alignment with recent standards and open, adaptive frameworks will be essential for sustainable, future-proof analytics infrastructures.

Recent and Emerging References: Special attention should be paid to major new initiatives such as WCH-BDP for clinical data integration [64, 94], cross-modal methods in urban and industrial analytics [6, 26, 92], and standardization projects like IBSI [1, 84] and radiomics harmonization frameworks [36]. As documented in recent surveys [8, 49, 53, 78, 79], the field is rapidly advancing toward more integrated, explainable, and resource-efficient solutions, while ongoing work continues to address interoperability, data diversity, and regulatory complexity.

11 Conclusions, Synthesis, and Outlook

This survey set out to achieve two principal objectives: (1) to provide a comprehensive synthesis of the most significant developments

and persistent challenges in the interdisciplinary intersection of AI domains, and (2) to foreground open research frontiers where integrative approaches are most needed. Our holistic overview emphasizes the need for unified conceptual frameworks and points out critical directions for ongoing and future research.

We have systematically examined how the integration of methods from machine learning, knowledge representation, and human-computer interaction has enabled substantial advances, while also identifying key obstacles—chiefly, methodological incompatibilities and the absence of common evaluation strategies. Going forward, fostering deeper collaboration across these domains will be essential for designing robust and interoperable AI systems.

To aid comparative analysis and synthesis, Table 23 provides a summary of major technical approaches surveyed in this work, highlighting their core methodologies, integration challenges, and evaluation mechanisms. This tabular overview is intended as a concise reference for both established and new researchers interested in cross-domain AI.

Summarizing our synthesis, unified frameworks that span AI subdomains remain an urgent research need. The emergence of early shared vocabularies and benchmarks has helped align community understanding, but more dynamic, extensible structures are required for ongoing integration. Clear specification of objectives, rigorous cross-domain evaluation, and the development of interoperable protocols will form the bedrock of the next generation of AI systems.

Based on the reviewed literature and our analysis, we distill the following explicit research questions and gaps that merit focused investigation:

How can standardized frameworks be established to support interoperability between emerging AI architectures across fields?

What strategies can ensure robust evaluation practices that reflect both domain-specific requirements and generalized performance expectations?

Which methodologies best facilitate transparency, interpretability, and human-in-the-loop collaboration—especially when integrating diverse AI technologies?

How can learned representations and reasoning processes be meaningfully shared and exchanged between distinct subdomains?

What socio-technical factors must be considered to ensure equitable, fair, and broadly beneficial outcomes, particularly for AI systems operating at disciplinary intersections?

Progress on these open questions will reinforce both theoretical insight and practical deployment of cross-disciplinary AI. Increased harmonization, through shared evaluation protocols and vocabulary, will empower future systems to address real-world challenges that span disciplinary boundaries.

We anticipate that continued synthesis and integration—guided by the comparative frameworks and open questions outlined here—will drive the next wave of transformative advances in AI.

Table 22: Comparative Overview of Secure, Scalable Analytical Ecosystem Approaches

Approach	Key Strengths	Principal Limitations	Representative Domains	Notable References
Federated Learning	Privacy-preserving model training; strong generalizability	Data heterogeneity, communication bottlenecks, update synchronization	Healthcare (medical imaging, predictive outcomes), IoT	[41, 64, 79, 94]
Blockchain Integration	Tamper-evident auditability, provenance assurance	System complexity, transaction latency, resource demands	Regulated healthcare, sensitive IoT, industrial systems	[1, 16, 35, 62, 65, 79, 91, 93]
Data Fusion Techniques	Scalable integration of heterogeneous modalities; improved analytics	Standardization, interoperability, and governance challenges	Energy, scientific research, industrial control, digital health	[6, 22, 26, 33, 53, 54, 92, 99, 102]
Standardized Platforms	End-to-end governance; efficient, reproducible workflows	Implementation complexity, evolving requirements, resource inequality	Hospital-wide health, EHR/omics research platforms	[1, 13, 27, 36, 47, 64, 84, 94, 95]
Explainable/Trustworthy AI	Transparency, regulatory compliance, future-proof design	High annotation overhead, deployment complexity	Medical AI, molecular imaging, precision medicine	[28, 34, 49, 53]

Table 23: Comparative overview of key cross-domain AI approaches, their methodologies, integration challenges, and evaluation mechanisms.

Approach	Core Methodologies	Integration Challenges	Evaluation Mechanisms
Symbolic-Connectionist Integration	Neural-symbolic models, Hybrid learning	Representation mismatch, Scalability	Task-specific, Interpretability metrics
Human-in-the-loop AI	Interactive learning, Feedback-driven adaptation	User alignment, Usability	Engagement rates, Satisfaction analysis
Knowledge Representation Interchange	Ontological mapping, Schema translation	Semantic consistency, Standardization	Benchmark tasks, Alignment accuracy
Cross-domain Transfer Learning	Domain adaptation, Meta-learning	Data heterogeneity, Generalization gap	Cross-task accuracy, Robustness tests
Interdisciplinary Benchmarking	Multi-faceted test suites, Shared tasks	Metric harmonization, Coverage	Composite scores, Standardized reporting

11.1 Synthesis of Advances Across Biomedical Signal Processing, Multimodal and Cross-Modal Analytics, Deep Learning, IoT/Real-Time Systems, Synthetic Data, and Explainable AI

The rapid convergence of advances in biomedical signal processing, multimodal and cross-modal analytics, deep learning, IoT-enabled real-time monitoring, synthetic data generation, and explainable artificial intelligence (XAI) has fundamentally reshaped the landscape of healthcare analytics and clinical operations [1–15, 17–39, 41–44, 47–49, 51–64, 66–72, 75–81, 83–88, 91–94, 96, 97, 99–102]. Signal processing remains the methodological foundation for precise feature extraction from biosignals such as EEG, ECG, sEMG, and medical imaging. Advanced signal processing methods, including rTV-gPDC for robust nonlinear causality analysis in multivariate physiological time series [81], MODWT-based multiresolution decomposition for seizure prediction [37], and dual-polynomial multiresolution techniques for myoelectric sEMG [80], are now routinely employed to improve signal integrity, attenuate noise and artifacts, and address nonlinearity within both clinical and IoT-driven contexts [15, 21, 37, 80, 81, 91]. Recent comparative studies also focus on optimizing time-frequency transformation methods (CWT vs. STFT) for ECG signal classification, revealing the superiority of certain wavelets in diagnostic performance and computational speed when feeding deep networks [86].

In parallel, deep learning architectures—notably convolutional neural networks (CNNs), transformers, and advanced encoder-decoder models—have established state-of-the-art benchmarks in medical data classification, detection, segmentation, and registration, frequently achieving diagnostic performance matching or surpassing human experts [1, 13, 42, 52, 53, 55, 58, 76, 85]. For example, 3D-UNet-based architectures with dense and residual blocks enable efficient and accurate coronary artery segmentation from CCTA [85], multiresolution CNNs support robust preictal/interictal EEG discrimination for seizure prediction [37], and dual-stream transformer frameworks (e.g., TransMatch) promote feature-level alignment for unsupervised deformable image registration [19]. The effectiveness of these models is significantly enhanced by sophisticated innovations in data fusion and multimodal learning.

Frameworks leveraging cooperative, late, early, and intermediate joint representation learning, as well as cross-modal adaptation and robust multi-view co-training, efficiently resolve challenges associated with high heterogeneity, complementarity, and missingness, thereby enabling robust domain adaptation and comprehensive modeling of both physiological and regulatory processes [2, 19, 23, 28, 30, 32, 36, 37, 41, 51, 62–64, 75, 79, 88, 92, 93, 97, 102]. Cutting-edge studies further highlight advances in model selection for informative modality subsets under cardinality constraints [32], the application of cross-modal loss functions and causality interventions [41, 62, 75], and multi-view fusion for robust seizure detection in EEG [2].

These methodological advances are accompanied by a transformation of healthcare delivery, largely attributable to the integration of IoT and real-time systems. Ubiquitous sensor networks, scalable analytics pipelines, and federated computation now support point-of-care diagnostics, home monitoring (such as wearable EEG and IMU for epilepsy detection in unsupervised environments) [18, 49, 50, 60, 78], and the optimization of operational workflows across the health system [72, 78, 94, 99]. Novel proportionate data analytics methods for medical data streams in IoT contexts [78] and hybrid recurrent modeling for menstrual health monitoring [49] showcase the capabilities of adaptive, real-time clinical monitoring.

Key methodological progress also stems from the introduction of privacy-preserving synthetic data generation techniques, such as generative adversarial networks (GANs), variational autoencoders (VAEs), and agent-based simulations. These approaches address limitations in data sharing, rare-event modeling, and regulatory compliance, while simultaneously necessitating the development of robust digital trust, audit, and chain-of-custody frameworks to mitigate risks relating to bias, re-identification, and provenance tracking [29, 57, 79, 81]. For example, digital chain-of-custody solutions and differential privacy mechanisms are proposed to secure synthetic data use cases in clinical contexts [29], and aligned whole-slide image (WSI) registration methods facilitate large-scale analysis of tumor heterogeneity while greatly reducing expert workload [57].

Additionally, the embedding of explainable AI mechanisms—most notably, SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME), and post-hoc saliency mapping—directly facilitates regulatory transparency, clinician trust, and eventual clinical adoption. Recent works demonstrate the practical benefits of these mechanisms: SHAP analysis for self-explaining diabetes diagnosis interfaces [20], XAI-augmented multi-view EEG seizure detection [2], and comprehensive post-hoc explainability frameworks for biomedical time series classification that resonate with clinician expectations [39]. Empirical evidence demonstrates tangible improvements in diagnosis, risk prediction, and operational support [2, 6, 20, 29, 39, 75, 79]. Collectively, these advances are propelling a systemic transformation from reactive, retrospective healthcare toward a proactive, data-driven, and adaptive paradigm.

11.2 Recap of IoT-Specific Strategies: Proportionate Data Analytics, HRLS-TS for Real-Time Monitoring, and Dynamic Error Detection

IoT-focused data analytics methodologies have matured considerably, enabling them to address the inherent complexity, noise, and heterogeneity of biomedical and operational data streams. Proportionate Data Analytics (PDA) exemplifies these advancements by providing an adaptive framework designed to distinguish between natural data variation and authentic anomalies within heterogeneous real-time signals. This distinction is crucial for maintaining an optimal balance between sensitivity (detecting true anomalies) and specificity (reducing false alarms), thus preventing system overload and ensuring reliable system performance [78]. PDA adopts regression-based statistical modeling to dynamically calibrate error detection thresholds, using error and variation decomposition across multiple, recurring time windows. This enables the system to maintain high identification and delivery ratios regardless of fluctuations in data quality, as demonstrated by its performance across key metrics such as accuracy, identification ratio, delivery, variation factor, and processing time.

Additionally, the Hybrid Recurrent Long Short-term based Tyrannosaurus Search (HRLS-TS) algorithm highlights the effectiveness of integrating deep learning with evolutionary computation for real-time, context-aware biomedical monitoring. For example, in menstrual physiology tracking, HRLS-TS leverages LSTM-based temporal modeling, enabling it to handle nonstationary data streams, while incorporating adaptive metaheuristic optimization to facilitate hyperparameter and feature selection. This hybrid approach achieves both high predictive accuracy and computational efficiency under the demanding requirements of streaming IoT healthcare applications [49].

Collectively, innovations such as PDA and HRLS-TS exemplify the ongoing evolution in IoT healthcare analytics, reflecting the field's growing ability to manage signal variability and noise while maintaining scalability and delivering actionable insights effectively for both clinical and operational decision-making.

11.3 Implications for Disease Detection, Personalized/Precision Medicine, and Clinical/Operational Workflow Integration

This survey aims to systematically review the integration of multimodal data, cross-modal analytics, and AI-driven approaches across disease detection, precision medicine, and healthcare operations, highlighting technological trends, methodological progress, and implementation challenges. In this synthesis, we evaluate how the convergence of the discussed breakthroughs is accelerating healthcare transformation while linking these implications to the survey's core objectives.

The convergence of technological and methodological breakthroughs discussed herein is accelerating healthcare transformation across disease detection, personalized and precision medicine, and workflow optimization. For disease detection, proactive analytics leveraging multimodal and cross-modal signals—including organ-specific liquid biopsies, multi-omic profiles, and continuous sensor-based physiological monitoring—facilitate earlier risk identification, often preceding the onset of overt symptoms or irreversible tissue pathology [10, 14, 52, 57, 79]. Notably, in domains such as infectious disease and oncology, rapid, scalable, and high-resolution stratification is now enabled by recent advances: for example, robust and scalable molecular diagnostics, digital pathology, and AI-driven imaging and sequencing analysis facilitate stratification and risk assessment [9, 11, 27, 30, 43, 52, 56, 63, 79, 83, 96]. These advancements are supported by methods that address challenges such as data fusion across imaging, sequence, and clinical data, with deep learning and self-supervised models (e.g., SPamCo [63]) or improved segmentation and risk prediction approaches (e.g., for brain tumors [9], renal carcinoma [56], and artery segmentation in CT angiography [85]) enhancing disease detection and prognosis.

Transitioning from detection to intervention, personalized and precision medicine initiatives are increasingly feasible as granular, multimodal patient phenotyping becomes the norm; integration across clinical, molecular, imaging, sensor, and environmental data streams enables sophisticated analytics pipelines capable of modeling intricate interactions and inferring causal relationships [23, 28, 36, 41, 42, 44, 52, 53, 55, 57, 92, 101, 102]. Research demonstrates that real-time digital interventions—including continuous glucose monitoring feedback systems and multi-source longitudinal tracking—lead to more tailored treatments, dynamic intervention adjustments, improved adherence, and measurable outcome improvements for chronic and complex conditions [8, 50, 81, 86]. These developments are well-aligned with the survey's objective of clarifying how multi-source data analytics can directly impact individualized patient trajectories.

A further central objective is the clinical and operational integration of these methodologies. Automated systems for segmentation, scheduling, triage, and capacity management—enabled by deep learning, time-series analytics, and explainable AI—reduce diagnostic bottlenecks, optimize resource allocation, and minimize delays in critical care pathways (e.g., cardiology, surgery, radiology) [3, 6, 22, 25, 38, 55, 85, 86]. For example, AI-driven segmentation methods in imaging (e.g., coronary artery segmentation [85]) and robust ECG signal classification [86] provide scalable automation in routine workflows. In parallel, federated and cloud-based

hospital data platforms are driving the interoperability and accessibility of analytics across multi-institutional settings, though continuing challenges in data harmonization and standardization persist [17, 77]. The ongoing adoption of explainable AI and integrated feedback systems is laying the groundwork for a more transparent, collaborative, and continuously learning healthcare environment.

In summary, these implications underscore the overarching goals of this survey: to map the landscape of methods enabling earlier and more accurate disease detection, support actionable and personalized decision-making, and enhance workflow efficiencies through integrated multimodal and cross-modal analytics. The field is rapidly evolving toward more holistic, patient-centered, and operationally efficient healthcare, though barriers in standardization, equity, and real-world implementation must still be addressed.

11.4 Future and Persistent Challenges: Standardization, Global Scalability, Heterogeneity, Explainability, Digital Equity, Privacy/Security, Responsible AI, and Collaborative Research

Despite significant progress, several persistent and systemic challenges must be addressed to translate these advances into large-scale, equitable, and clinically impactful solutions. Chief among these is the need for rigorous standardization—both in feature definition (for example, through initiatives such as the Image Biomarker Standardization Initiative and advances in radiomics/texture analysis) and in clinical reporting. Even subtle inconsistencies in data computation or documentation can undermine reproducibility, limit interoperability, and impede clinical adoption [1–39, 41–44, 46–49, 51–64, 66–72, 75–81, 83–88, 91–94, 96, 97, 99, 100, 102].

Additional barriers include:

Lack of universal data standards: Inconsistent structure, labeling, and exchange protocols impede interoperability and multi-center collaborations [17, 68, 77].

Global scalability and heterogeneity: State-of-the-art AI and signal processing models often demonstrate reduced generalizability outside well-annotated, homogeneous cohorts, calling for improved domain adaptation, quality-aware fusion, and dynamic modality selection [32, 57, 71, 97].

Explainability: Progress toward actionable and domain-relevant explanations is critical for supporting clinical accountability, user trust, and regulatory acceptance [6, 20, 75, 79].

Digital equity and privacy/security: Variability in access to analytics, infrastructure, and digital literacy—both inter- and intra-regionally—risks entrenching health disparities. Concurrently, the transition to granular, continuous, interoperable data streams heightens privacy and security risks, necessitating progress in cryptography, differential privacy, and auditable governance [39, 68, 79, 81, 81, 86].

Responsible AI and collaborative research: Ensuring responsible, bias-mitigated AI deployment requires regulatory harmonization and inclusive frameworks that bring together multidisciplinary stakeholders across research, clinical, technical, regulatory, and patient spheres [1–39, 41–44, 46–49, 51–64, 66–72, 75–81, 83–88, 91–94, 96, 97, 99, 100, 102].

To provide a clear overview, the following table summarizes the persistent challenges along with their identified impacts and key needs:

In summary, the synthesis of progress across biomedical signal processing, multimodal and cross-modal analytics, deep learning, IoT/real-time systems, synthetic data methodologies, and explainable AI marks a decisive step toward more precise, predictive, and participatory healthcare. Nevertheless, realization of this promise will depend upon concerted and sustained efforts to standardize, validate, explain, secure, and equitably distribute these innovations at scale—always guided by ethical, social, and clinical imperatives.

References

[1] O. Aerts, S. W. Griffiths, M. W. A. van Timmeren, and et al. 2020. The Image Biomarker Standardisation Initiative: Standardized Quantitative Radiomics for High Throughput Image-based Phenotyping. *Radiology* 295, 2 (2020), 328–338. doi:10.1148/radiol.2020191145

[2] Ijaz Ahmad, Zhenzhen Liu, Lin Li, Inam Ullah, Sunday Timothy Aboyeji, Xin Wang, Oluwarotimi W. Samuel, Guanglin Li, Yuan Tao, Yan Chen, and Shixiong Chen. 2024. Robust Epileptic Seizure Detection Based on Biomedical Signals Using an Advanced Multi-View Deep Feature Learning Approach. *IEEE Journal of Biomedical and Health Informatics* 28, 10 (2024), To appear. doi:10.1109/JBHI.2024.3350919

[3] R. Ahmadirouhani, M. H. Karimpour, B. Rahimi, A. Malekzadeh-Shafaroudi, A. Beiranvand Pour, and B. Pradhan. 2018. Integration of SPOT-5 and ASTER satellite data for structural tracing and hydrothermal alteration mineral mapping: implications for Cu–Au prospecting. *International Journal of Image and Data Fusion* 9, 3 (2018), 239–253. doi:10.1080/19479832.2018.1469548

[4] C. M. Alonso, M. Llop, C. Sargas, L. Pedrola, J. Panadero, D. Hervás, J. Cervera, E. Such, M. Ibáñez, R. Ayala, J. Martínez-López, E. Onecha, I. de Juan, S. Palanca, D. Martínez-Cuadrón, R. Rodríguez-Veiga, B. Boluda, P. Montesinos, G. Sanz, M. A. Sanz, and E. Barragán. 2019. Clinical Utility of a Next-Generation Sequencing Panel for Acute Myeloid Leukemia Diagnostics. *The Journal of Molecular Diagnostics* 21, 2 (2019), 228–240. doi:10.1016/j.jmoldx.2018.09.009

[5] M. Antillon, M. McFarland, J. Ndung'u, E. Sutherland, M.E. Chappuis, S. Snijders, A. Hope, W.J. de Vlas, and E.M. Castaño. 2020. Cost-effectiveness of sleeping sickness elimination campaigns in the Democratic Republic of Congo: a transmission model-based economic evaluation. *medRxiv* (2020). doi:10.1101/2020.08.25.20181982

[6] R. J. Appleton, B. C. Barnes, and A. Strachan. 2025. Data Fusion of Deep Learned Molecular Embeddings for Property Prediction. *arXiv preprint arXiv:2504.07297* (2025). https://arxiv.org/abs/2504.07297

[7] F. Atilla, B. Klomberg, B. Cardoso, and N. Cohn. 2023. Background check: cross-cultural differences in the spatial context of comic scenes. *Multimodal Communication* 12, 3 (2023), 179–189. https://www.degruyter.com/document/doi/10.1515/mc-2023-2010/html

[8] Mohammed Badawy, Nagy Ramadan, and Hesham Ahmed Hefny. 2023. Healthcare predictive analytics using machine learning and deep learning techniques: a survey. *Journal of Electrical Systems and Information Technology* 10 (2023). doi:10.1186/s43067-023-00108-y

[9] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, R. Shinohara, et al. 2019. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *Medical Image Analysis* 55 (2019), 80–91. doi:10.1016/j.media.2019.04.015

[10] W. Barker, W. Chang, J. Everson, M. Gabriel, V. Patel, C. Richwine, and C. Strawley. 2024. The Evolution of Health Information Technology for Enhanced Patient-Centric Care in the United States: Data-Driven Descriptive Study. *Journal of Medical Internet Research* 26, 1 (2024), e59791. doi:10.2196/59791

[11] J. A. Bateman and F. Schmidt-Borcherding. 2018. The Communicative Effectiveness of Education Videos: Towards an Empirically-Motivated Multimodal Account. *Multimodal Technologies and Interaction* 2, 3 (2018), 59. https://www.mdpi.com/2414-4088/2/3/59

[12] J. A. Bateman and C.-I. Tseng. 2023. Multimodal discourse analysis as a method for revealing narrative strategies in news videos. *Multimodal Communication* 12, 3 (2023), 261–285. doi:10.1515/mc-2023-2017

[13] D. Benielli, H. Berry, and J. Morlier. 2022. scikit-multimodallearn: A Python Toolbox for Multimodal Machine Learning. *Journal of Machine Learning Research* 23, 279 (2022), 1–6. https://www.jmlr.org/papers/volume23/21-0791/21-0791.pdf

[14] Béatrice Biancardi, Maurizio Mancini, Brian Ravenet, and Giovanna Varni. 2024. Modelling the 'transactive memory system' in multimodal multiparty interactions. *Journal on Multimodal User Interfaces* 18, 1 (2024), 103–117.

Table 24: Persistent Challenges and Required Interventions for Scaling Biomedical AI in Healthcare

Challenge	Impact	Key Needs and Solutions
Standardization	Limits reproducibility, interoperability, clinical trust	Universal data formats, reporting standards, feature harmonization (e.g., Image Biomarker Standardization Initiative)
Global Scalability and Heterogeneity	Reduces accuracy in diverse/underrepresented groups	Domain adaptation, quality-aware fusion, data augmentation for imbalanced datasets
Explainability	Slows clinical adoption, impedes trust and accountability	Domain-relevant, actionable explanations (e.g., SHAP, LIME); regulatory guidance on XAI
Digital Equity and Privacy/Security	Exacerbates health disparities, threatens confidentiality	Policy-driven infrastructure investments, digital literacy programs, advanced privacy and security technologies
Responsible AI and Collaborative Research	Risks bias, misdeployment, ethical lapses	Inclusive stakeholder collaboration, rigorous validation, transparent and ethical governance

doi:10.1007/s12193-023-00399-w

[15] F. Byrne. 2024. A diffractive approach to multimodal transcription: Materialising entanglements between humans and non-humans. *Multimodality & Society* (2024). doi:10.1177/26349795241291369

[16] N. Bürgisser, E. Chalot, S. Mehouachi, C. P. Buclin, K. Lauper, D. S. Courvoisier, and D. Mongin. 2024. Large language models for accurate disease detection in electronic health records. *medRxiv* (2024). https://www.medrxiv.org/content/10.1101/2024.07.27.24311106v1

[17] Kavyashree C., H. S. Vimala, and Shreyas J. 2024. A systematic review of artificial intelligence techniques for oral cancer detection. *Healthcare Analytics* 7 (2024), 100304. doi:10.1016/j.health.2024.100304

[18] Christos Chatzichristos, Lauren Swinnen, Jaiver Macea, Miguel Bhagubai, Wim Van Paesschen, and Maarten De Vos. 2022. Multimodal detection of typical absence seizures in home environment with wearable electrodes. *Frontiers in Signal Processing* 2 (2022), 1014700. doi:10.3389/frsip.2022.1014700

[19] Z. Chen, Y. Zheng, and J. C. Gee. 2024. TransMatch: A Transformer-Based Multilevel Dual-Stream Feature Matching Network for Unsupervised Deformable Image Registration. *IEEE Transactions on Medical Imaging* 43, 1 (2024), 15–27. doi:10.1109/TMI.2023.3324623

[20] G. Dharmarathne, T. N. Jayasinghe, M. Bogahawaththa, D. Meddage, and U. Rathnayake. 2024. A novel machine learning approach for diagnosing diabetes with a self-explainable interface. *Healthcare Analytics* 7 (2024), 100301. doi:10.1016/j.health.2024.100301

[21] Hamdi Dibeklioğlu, Elif Surer, Albert Ali Salah, and Thierry Dutoit. 2021. Behavior and usability analysis for multimodal user interfaces. *Journal on Multimodal User Interfaces* 15 (2021), 335–336. doi:10.1007/s12193-021-00372-0

[22] A. Diez-Olivan, J. Del Ser, D. Galar, and B. Sierra. 2019. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Information Fusion* 50 (Jan. 2019), 92–111. doi:10.1016/j.inffus.2018.10.005

[23] D. Ding, S. Li, B. Narasimhan, and R. Tibshirani. 2022. Cooperative learning for multiview analysis. *Nature Machine Intelligence* 4 (2022), 263–271. https://www.nature.com/articles/s42256-022-00497-6

[24] I. J. Borges do Nascimento, M. S. Marcolino, H. M. Abdulazeem, I. Weerasekara, N. Azzopardi-Muscat, M. A. Gonçalves, and D. Novillo-Ortiz. 2021. Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies. *Journal of Medical Internet Research* 23, 4 (2021), e27275. doi:10.2196/27275

[25] P. Dutta. 2018. An uncertainty measure and fusion rule for conflict evidences of big data via Dempster–Shafer theory. *International Journal of Image and Data Fusion* 9, 2 (2018), 187–201. https://www.tandfonline.com/doi/full/10.1080/19479832.2017.1391336

[26] W. Feng, Q. Li, C. Wang, and J. s. Fan. 2025. Data Fusion for Full-Range Response Reconstruction via Diffusion Models. *arXiv preprint arXiv:2502.00795* (2025). https://arxiv.org/abs/2502.00795

[27] S. Flynn, A. Haenel, F. Coughlan, S. Crilly, J. A. Leipsic, and J. D. Dodd. 2024. Cardiac CT, MRI, and PET in 2023: Exploration of Key Articles across Imaging and Multidisciplinary Journals. *Radiology* 313, 3 (2024), e240975. doi:10.1148/radiol.240975

[28] F. Fraioli, N. Albert, R. Boellaard, I. Boscolo Galazzo, M. Brendel, I. Buvat, M. Castellaro, D. Cecchin, P. Aguiar Fernandez, E. Guedj, A. Hammers, Z. Kaplar, S. Morbelli, L. Papp, K. Shi, N. Tolboom, T. Traub-Weidinger, A. Verger, D. Van Weehaeghe, I. Yakushev, and H. Barthel. 2024. Perspectives of the European Association of Nuclear Medicine on the role of artificial intelligence (AI) in molecular brain imaging. *European Journal of Nuclear Medicine and Molecular Imaging* 51 (2024), 1007–1011. doi:10.1007/s00259-023-06553-1

[29] M. Giuffrè et al. 2023. Harnessing the power of synthetic data in healthcare. *npj Digital Medicine* 6 (2023), Article No. 198. doi:10.1038/s41746-023-00927-3

[30] P. Goyal, S. Sahu, S. Ghosh, and C. Lee. 2020. Cross-modal Learning for Multimodal Video Categorization. *arXiv preprint arXiv:2003.03501* (2020). https://arxiv.org/abs/2003.03501

[31] M. Habib, Z. Wang, S. Qiu, H. Liu, J. Qi, P. Kumar, and W. Chen. 2022. Machine Learning Based Healthcare System for Investigating the Association Between Depression and Quality of Life. *IEEE Journal of Biomedical and Health Informatics* 26, 6 (2022), 2795–2805. doi:10.1109/JBHI.2022.3152053

[32] Y. He, L. Zheng, X. Shi, and T. Huang. 2024. Efficient Modality Selection in Multimodal Learning. *Journal of Machine Learning Research* 25, 61 (2024), 1–53. https://www.jmlr.org/papers/volume25/23-0439/23-0439.pdf

[33] Y. Himeur, A. Ghanem, F. Alsalemi, R. Alhameed, F. Bensaali, and A. Amira. 2020. Data fusion strategies for energy efficiency in buildings: Overview, taxonomy, and future research directions. *Information Fusion* 64 (Oct. 2020), 99–120. https://linkinghub.elsevier.com/retrieve/pii/S1566253520303158

[34] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller. 2022. Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Information Fusion* 79 (Nov. 2022), 263–278. https://linkinghub.elsevier.com/retrieve/pii/S1566253521002050

[35] A. Hope, C. F. Dietrich, L. Q. Zhou, D. L. McFarlane, R. J. Gillies, and H. Hricak. 2016. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* 278, 2 (2016), 563–577. https://pubs.rsna.org/doi/10.1148/radiol.2015151169

[36] P. Hu, Z. Huang, D. Peng, X. Wang, and X. Peng. 2023. Cross-modal retrieval with partially mismatched pairs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 8 (2023), 9595–9610. https://ieeexplore.ieee.org/document/10050111/

[37] Ali K. Ibrahim, Hanqi Zhuang, Emmanuelle Tognoli, Ali Muhamed Ali, and Nurgun Erdol. 2023. Epileptic seizure prediction based on multiresolution convolutional neural networks. *Frontiers in Signal Processing* 2 (2023), 1175305. doi:10.3389/frsip.2023.1175305

[38] S. Iino, R. Ito, K. Doi, T. Imaizumi, S. Hikosaka, and T. Y. Saito. 2018. CNN-based generation of high-accuracy urban distribution maps utilising SAR satellite imagery for short-term change monitoring. *International Journal of Image and Data Fusion* 9, 4 (2018), 298–312. doi:10.1080/19479832.2018.1491897

[39] P. Ivaturi, D. Chen, T. M. Talbert, Z. Jiang, W. Xu, G. Quer, M. Gadaleta, D. Jha, J. C. Lin, S. Santini, and P. Bonato. 2021. A Comprehensive Explanation Framework for Biomedical Time Series Classification. *IEEE Journal of Biomedical and Health Informatics* 25, 7 (2021), 2398–2408. doi:10.1109/JBHI.2021.3060997

[40] B. R. Jackson, C. Tarr, E. Strain, K. A. Jackson, A. Conrad, H. Carleton, L. S. Katz, S. Stroika, L. H. Gould, R. K. Mody, B. J. Silk, J. Beal, Y. Chen, R. Timme, M. Doyle, A. Fields, M. Wise, G. Tillman, S. Defibaugh-Chavez, Z. Kucerova, A. Sabol, K. Roache, E. Trees, M. Simmons, J. Wasilenko, K. Kubota, H. Pouseele, W. Klimke, J. Besser, E. Brown, M. Allard, and P. Gerner-Smidt. 2016. Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation. *Clinical Infectious Diseases* 63, 3 (2016), 380–386. doi:10.1093/cid/ciw242

[41] M. Jaritz, T.-H. Vu, R. De Charette, É. Wirbel, and P. Pérez. 2022. Cross-modal learning for domain adaptation in 3D semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 2 (2022), 1533–1544. https://ieeexplore.ieee.org/document/9737217/

[42] L. J. Jennings, M. E. Arcila, C. Corless, S. Kamel-Reid, I. M. Lubin, J. Pfeifer, R. L. Temple-Smolkin, K. V. Voelkerding, and M. N. Nikiforova. 2017. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *The Journal of Molecular Diagnostics* 19, 3 (2017), 341–365. doi:10.1016/j.jmoldx.2017.01.011

[43] S. Kjær Jensen and B. Schirrmacher. 2024. Stronger together: Moving towards a combined multimodal and intermedial model. *Multimodality & Society* (2024). doi:10.1177/26349795241259606

[44] Yushan Jiang, Kanghui Ning, Zijie Pan, Xuyang Shen, Jingchao Ni, Wenchao Yu, Anderson Schneider, Haifeng Chen, Yuriy Nevmyvaka, and Dongjin Song. 2025. Multi-modal Time Series Analysis: A Tutorial and Survey. *arXiv preprint arXiv:2503.13709* (2025). https://arxiv.org/abs/2503.13709

[45] L. M. R. Kasbergen, E. de Bruin, F. Chandler, A. J. Menting, P. A. M. Kouwijzer, L. Schwanitz, V. Klaus, K. H. Adan, S. Markota, R. G. P. Alblas, et al. 2024. Multi-antigen serology and a diagnostic algorithm for the detection of arbovirus infections as novel tools for arbovirus preparedness in southeast Europe: a prospective observational study. *The Lancet Infectious Diseases* 24, 6 (2024), 689–698. doi:10.1016/S1473-3099(24)00654-6

[46] M. V. Kiang, E. T. Chin, B. Q. Huynh, L. A. C. Chapman, I. Rodríguez-Barraquer, B. Greenhouse, G. W. Rutherford, K. Bibbins-Domingo, D. Havlir, S. Basu, and N. C. Lo. 2021. Routine asymptomatic testing strategies for airline travel during the COVID-19 pandemic: a simulation study. *The Lancet Infectious Diseases* 21, 7 (2021), 929–938. doi:10.1016/S1473-3099(21)00134-1

[47] E. M. Kudalkar, N. A. M. Almontashiri, C. Huang, B. Anekella, M. Bowser, E. Hynes, R. Garlick, and B. H. Funke. 2016. Multiplexed Reference Materials as Controls for Diagnostic Next-Generation Sequencing: A Pilot Investigating Applications for Hypertrophic Cardiomyopathy. *The Journal of Molecular*

Diagnostics 18, 6 (2016), 882–889. doi:10.1016/j.jmoldx.2016.07.005

[48] A. Kumar, P. Nandakishore, A. J. Gordon, E. Baum, J. Madhok, Y. Duanmu, and J. Kugler. 2025. Creation of an Open-Access Lung Ultrasound Image Database For Deep Learning and Neural Network Applications. *medRxiv* (2025). doi:10.1101/2025.05.09.25327337

[49] P. M. Kumar, U. P. Rao, P. Vijayakumar, N. Chilamkurti, D. Q. Phung, and N. Kumaravel. 2022. Clouds Proportionate Medical Data Stream Analytics for Internet of Things-Based Healthcare Systems. *IEEE Journal of Biomedical and Health Informatics* 26, 3 (2022), 973–982. https://ieeexplore.ieee.org/document/9520255/

[50] J. Lamb. 2018. To Boldly Go: Feedback as Digital, Multimodal Dialogue. *Multimodal Technologies and Interaction* 2, 3 (2018), 49. https://www.mdpi.com/2414-4088/2/3/49

[51] C. Lauri, A. Signore, A.W.J.M. Glaudemans, G. Treglia, O. Gheysens, R.H.J.A. Slart, R. Iezzi, N.H.J. Prakken, E.S. Debus, S. Honig, A. Lejay, and N. Chakfé. 2022. Evidence-based guideline of the European Association of Nuclear Medicine (EANM) on imaging infection in vascular grafts. *European Journal of Nuclear Medicine and Molecular Imaging* 49 (2022), 3430–3451. doi:10.1007/s00259-022-05769-x

[52] M. B. Lawson, W. Zhu, D. L. Miglioretti, T. Onega, L. M. Henderson, G. H. Rauscher, K. Kerlikowske, B. L. Sprague, E. J. A. Bowles, E. S. O'Meara, A. N. A. Tosteson, R. M. diFlorio Alexander, R. A. Hubbard, J. M. Lee, and C. I. Lee. 2025. Disparities in Standard-of-Care, Advanced, and Same-Day Diagnostic Services among Patients with Abnormal Screening Mammography. *Radiology* 314, 2 (2025), e241673. doi:10.1148/radiol.241673

[53] H. Laçi, K. Sevrani, and S. Iqbal. 2025. Deep learning approaches for classification tasks in medical X-ray, MRI, and ultrasound images: a scoping review. *BMC Medical Imaging* 25, 1 (2025), 156. https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-024-01122-0

[54] Hyeonhoon Lee, Hyun-Lim Yang, Hyuk Goo Ryu, Chiwon Jung, Yong Joon Cho, Seungbeom Yoon, Hyun-Kyung Yoon, and Hee-Chang Lee. 2023. Real-time machine learning model to predict in-hospital cardiac arrest using heart rate variability in ICU. *npj Digital Medicine* 6 (2023), Article no. 215. doi:10.1038/s41746-023-00960-2

[55] J. Li, P. Zhang, T. Wang, L. Zhu, R. Liu, X. Yang, K. Wang, D. Shen, and B. Sheng. 2024. DSMT-Net: Dual Self-Supervised Multi-Operator Transformation for Multi-Source Endoscopic Ultrasound Diagnosis. *IEEE Transactions on Medical Imaging* 43, 1 (2024), 64–75. doi:10.1109/TMI.2023.3324373

[56] X. Li, Y. Guo, S. Huang, F. Wang, C. Dai, J. Zhou, and D. Lin. 2025. A CT-based intratumoral and peritumoral radiomics nomogram for postoperative recurrence risk stratification in localized clear cell renal cell carcinoma. *BMC Medical Imaging* 25, 1 (2025), 167. doi:10.1186/s12880-024-01133-x

[57] Xiaoxiao Li, Liye Mei, Mengping Long, Jin Huang, Wei Yang, Xin Hao, Yi qang Liu, Jinxuan Hou, Yu Xu, Fuling Zhou, Dupei Wang, Jianghua Wu, Taobo Hu, and Cheng Lei. 2025. Accurately matching serial whole slide images for tumor heterogeneity analysis. *Biomedical Signal Processing and Control* 87 (2025), 106825. doi:10.1016/j.bspc.2024.106825

[58] Z. Lin, S. Yu, Z. Kuang, D. Pathak, and D. Ramanan. 2023. Multimodality Helps Unimodality: Cross-Modal Few-Shot Learning with Multimodal Models. *arXiv preprint arXiv:2301.06267* (2023). https://arxiv.org/abs/2301.06267

[59] F. H. Linden, M. Boesen, E. Høegh-schmidt, J. U. Nybing, R. B. Mainz Hansen, L. H. Persson Kopp, M. N. Thomsen, and M. W. Brejnebøl. 2025. External Validation of an Automated Segmentation Tool for Abdominal Fat Tissue using DIXON MRI: Data from the CutDM trial. *medRxiv* (2025). doi:10.1101/2025.05.14.25327583

[60] D. Lindenberg. 2023. Multimodal meaning-making in student presentations: the impact of explicit feedback in a German as a foreign language classroom. *Multimodal Communication* 12, 3 (2023), 191–206. doi:10.1515/mc-2023-2011

[61] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis* 42 (2017), 60–88. doi:10.1016/j.media.2017.07.005

[62] Y. Liu, G. Li, and L. Lin. 2023. Cross-modal causal relational reasoning for event-level visual question answering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 10 (2023), 11624–11641. https://ieeexplore.ieee.org/document/10146482/

[63] F. Ma, C. Liu, C. Aggarwal, J. Gao, and J. Han. 2020. Self-paced Multi-view Co-training. *Journal of Machine Learning Research* 21, 182 (2020), 1–31. https://www.jmlr.org/papers/volume21/18-794/18-794.pdf

[64] S. J. Mataraso, C. A. Espinosa, D. Seong, S. M. Reincke, E. Berson, J. D. Reiss, Y. Kim, M. Ghanem, C.-H. Shu, T. James, Y. Tan, S. Shome, I. A. Stelzer, D. Feyaerts, R. J. Wong, G. M. Shaw, M. S. Angst, B. Gaudilliere, D. K. Stevenson, and N. Aghaeepour. 2025. A machine learning approach to leveraging electronic health records for enhanced omics analysis. *Nature Machine Intelligence* 7 (2025), 293–306. doi:10.1038/s42256-024-00974-9

[65] A. Matov. 2024. Urinary Biomarkers for Disease Detection. *medRxiv* (2024). https://www.medrxiv.org/content/10.1101/2024.07.30.24311186v1

[66] Candy Olivia Mawalim, Shogo Okada, Yukiko I. Nakano, and Masashi Unoki. 2023. Personality trait estimation in group discussions using multimodal analysis and speaker embedding. *Journal on Multimodal User Interfaces* 17 (2023), 47–63. doi:10.1007/s12193-023-00401-0

[67] Camille Mazzara, Pietro Sorrentino, Irene Altuna, Roma Maghsoudlou, Fabio Ferrarelli, Diego Sona, Luca Cocchi, Alessandro Gozzi, Carlo Miniussi, Silvia De Santis, Francesca Pistoia, Kenji Kansaku, Pau Erola, Gregoire Danaila, Carlo Cavedini, Barry A. Giesbrecht, Gordon Sajda, Rodolphe Jenatton, and Emmanuel Daucé. 2024. Mapping brain lesions to conduction delays: the next step for personalized brain models in Multiple Sclerosis. *medRxiv* (2024). doi:10.1101/2024.07.10.24310150 Preprint.

[68] N. Mehta, Z. Pandit, and S. Shukla. 2019. Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. *Journal of Biomedical Informatics* 100 (2019), Article 103311. doi:10.1016/J.JBI.2019.103311

[69] F. Meng, T. Zhang, Y. Pan, X. Kan, Y. Xia, M. Xu, J. Cai, F. Liu, and Y. Ge. 2025. A deep learning algorithm for automated adrenal gland segmentation on non-contrast CT images. *BMC Medical Imaging* 25, 1 (2025), 142. doi:10.1186/s12880-025-01682-5

[70] S. Mohammed, T.-H. Kim, R.-S. Chang, and C. Ramos. 2022. Guest Editorial: Data Analytics for Public Health Care. *IEEE Journal of Biomedical and Health Informatics* 26, 4 (2022), 1409–1410. https://ieeexplore.ieee.org/document/9757035/

[71] T. Nguyen-Ky, H. D. Tuan, A. V. Savkin, M. N. Do, and N. T. T. Van. 2021. Real-Time EEG Signal Classification for Monitoring and Predicting the Transition Between Different Anaesthetic States. *IEEE Transactions on Biomedical Engineering* 68, 5 (2021), 1450–1458. doi:10.1109/TBME.2021.3053019

[72] M. Nikolova-Simons, R. Keldermann, Y. Peters, W. Compagner, L. Montenij, Y. de Jong, and R. A. Bouwman. 2023. Predictive analytics for cardio-thoracic surgery duration as a stepstone towards data-driven capacity management. *npj Digital Medicine* 6 (2023). doi:10.1038/s41746-023-00938-0

[73] I. N. Okeke and C. Ihekweazu. 2021. The importance of molecular diagnostics for infectious diseases in low-resource settings. *Nature Reviews Microbiology* 19 (2021), 547–548. doi:10.1038/s41579-021-00598-5

[74] A. Pierangeli, O. Turriziani, M. Fracella, R. Campagna, F. Frasca, A. D'Auria, C. Scagnolari, P. Roberto, L. Cinti, G. D'Ettorre, G. Ceccarelli, L. Petrarca, R. Nenna, F. Midulla, G. Galardo, and G. Antonelli. 2025. The added value of diagnostics to characterize age-specific patterns of respiratory viral infections and coinfections and to detect emerging threats. *BMC Infectious Diseases* 25 (2025), 404. doi:10.1186/s12879-025-10693-0

[75] Mirco Planamente, Chiara Plizzari, Simone Alberto Peirone, Barbara Caputo, and Andrea Bottino. 2024. Relative Norm Alignment for Tackling Domain Shift in Deep Multi-modal Classification. *International Journal of Computer Vision* 132 (2024), 2618–2638. doi:10.1007/s11263-024-01998-9

[76] E. H. P. Pooch, G. Agrotis, L. Cai, M. Emberton, T. T. Shah, H. U. Ahmed, R. G. H. Beets-Tan, S. Benson, T. Janssen, and I. G. Schoots. 2025. Semi-Supervised Learning in Prostate MRI Tumor Segmentation Approaches Fully-Supervised Performance on External Validation. *medRxiv* (2025). doi:10.1101/2025.05.13.25327456

[77] R. A. Rasul, P. K. Saha, D. Bala, S. Karim, Md. I. Abdullah, and B. P. Saha. 2024. An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder. *Healthcare Analytics* 7 (2024), 100293. doi:10.1016/j.health.2024.100293

[78] M. P. Recht, L. M. White, J. Fritz, and D. L. Resnick. 2023. Advances in Musculoskeletal Imaging: Recent Developments and Predictions for the Future. *Radiology* 308, 2 (2023), e230615. doi:10.1148/radiol.230615

[79] H. R. Roth, X. Zhuang, D. Shen, and C. Shen. 2022. Federated learning for medical image analysis. *Medical Image Analysis* 75 (2022), 102298. doi:10.1016/j.media.2021.102298

[80] O. W. Samuel, Y. Z. Kulwa, G. Li, R. N. Khushaba, and S. S. Ge. 2022. Multiresolution Dual-Polynomial Decomposition Approach for Optimized Characterization of Motor Intent in Myoelectric Control Systems. *IEEE Transactions on Biomedical Engineering* 69, 10 (2022), 3122–3133. doi:10.1109/tbme.2022.3221799

[81] T. Schäck, M. Muma, M. Feng, C. Guan, and A. M. Zoubir. 2018. Robust Nonlinear Causality Analysis of Nonstationary Multivariate Physiological Time Series. *IEEE Transactions on Biomedical Engineering* 65, 6 (2018), 1213–1225. https://ieeexplore.ieee.org/document/7934390

[82] Gizachew Mulu Setegn and Belayneh Endalamaw Dejene. 2025. Explainable AI for Symptom-Based Detection of Monkeypox: a machine learning approach. *BMC Infectious Diseases* 25 (2025). doi:10.1186/s12879-025-10738-4

[83] Chengdong Shi, Kai Yu, Yu Hu, Yuantao Wang, Fan Bu, Ji Lu, and Weigang Wang. 2024. Diagnostic Efficacy of Various Imaging Modalities Across Different Stages of Prostate Cancer: A Network Meta-Analysis of Diagnostic Studies. *medRxiv* (2024). doi:10.1101/2024.09.28.24314285

[84] F. Sollini, R. Antunovic, A. Chiti, and H. J. W. L. Aerts. 2021. PET Radiomics in NSCLC: State of the Art and a Proposal for Harmonization of Methodology. *Radiology* 298, 2 (2021), 548–560. doi:10.1148/radiol.2021200616

[85] A. Song, L. Xu, L. Wang, B. Wang, X. Yang, B. Xu, B. Yang, and S. E. Greenwald. 2022. Automatic coronary artery segmentation of CCTA images with an efficient

feature-fusion-and-rectification 3D-UNet. *IEEE Journal of Biomedical and Health Informatics* 26, 8 (2022), 4044–4055. https://ieeexplore.ieee.org/document/9405480

[86] Min-Seo Song and Seung-Bo Lee. 2024. Comparative study of time-frequency transformation methods for ECG signal classification. *Frontiers in Signal Processing* 4 (2024). doi:10.3389/frsip.2024.1322334

[87] S. S. Srinivas and V. Runkana. 2024. Cross-Modal Learning for Chemistry Property Prediction: Large Language Models Meet Graph Machine Learning. *arXiv preprint arXiv:2408.14964* (2024). https://arxiv.org/abs/2408.14964

[88] S. R. Stahlschmidt, B. Ulfenborg, and J. Synnergren. 2023. Multimodal deep learning for biomedical data fusion: a review. *Nature Machine Intelligence* 5 (2023), 351–362. https://www.nature.com/articles/s42256-023-00629-w

[89] A. J. Sundermann, J. Chen, P. Kumar, A. M. Ayres, S. T. Cho, C. Ezeonwuka, M. P. Griffith, J. K. Miller, M. M. Mustapha, A. W. Pasculle, M. I. Saul, K. A. Shutt, V. Srinivasa, K. Waggle, D. J. Snyder, V. S. Cooper, D. Van Tyne, G. M. Snyder, J. W. Marsh, A. Dubrawski, M. S. Roberts, and L. H. Harrison. 2022. Whole-Genome Sequencing Surveillance and Machine Learning of the Electronic Health Record for Enhanced Healthcare Outbreak Detection. *Clinical Infectious Diseases* 75, 3 (2022), 476–482. doi:10.1093/cid/ciab946

[90] A. J. Sundermann, P. Kumar, M. P. Griffith, K. D. Waggle, V. Rangachar Srinivasa, N. Raabe, E. G. Mills, H. Coyle, D. Ereifej, H. M. Creager, A. Ayres, D. Van Tyne, L. L. Pless, G. M. Snyder, M. Roberts, and L. H. Harrison. 2025. Real-Time Genomic Surveillance for Enhanced Healthcare Outbreak Detection and Control: Clinical and Economic Impact. *Clinical Infectious Diseases* (2025). doi:10.1093/cid/ciaf216

[91] D. van Kampen and J. Wildfeuer. 2024. Cohesion in physical movie releases: developing a framework for the multimodal analysis of complex three-dimensional packaging. *Multimodality & Society* (2024). doi:10.1177/26349795241306993

[92] A. Vobecky, D. Hurych, O. Siméoni, S. Gidaris, A. Bursuc, P. Pérez, and J. Sivic. 2025. Unsupervised Semantic Segmentation of Urban Scenes via Cross-Modal Distillation. *International Journal of Computer Vision* 133, 6 (2025), 3519–3541. doi:10.1007/s11263-024-02320-3

[93] Jiaqi Wang, Hanqi Jiang, Yiheng Liu, Chong Ma, Xu Zhang, Yi Pan, Mengyuan Liu, Peiran Gu, Sichen Xia, Wenjun Li, Yutong Zhang, Zihao Wu, Zhengliang Liu, Tianyang Zhong, Bao Ge, Tuo Zhang, Ning Qiang, Xintao Hu, Xi Jiang, Xin Zhang, Wei Zhang, Dinggang Shen, Tianming Liu, and Shu Zhang. 2024. A Comprehensive Review of Multimodal Large Language Models: Performance and Challenges Across Different Tasks. *arXiv preprint arXiv:2408.01319* (2024). https://arxiv.org/abs/2408.01319

[94] M. Wang, S. Li, T. Zheng, N. Li, Q. Shi, X. Zhuo, R. Ding, and Y. Huang. 2022. Big Data Health Care Platform With Multisource Heterogeneous Data Integration and Massive High-Dimensional Data Governance for Large Hospitals: Design, Development, and Application. *JMIR Medical Informatics* 10, 4 (2022), e36481. doi:10.2196/36481

[95] T. Weitzel, C. Pérez, D. Tapia, P. Legarraga, and L. Porte. 2021. SARS-CoV-2 rapid antigen detection tests. *The Lancet Infectious Diseases* 21, 8 (2021), 1067–1068. doi:10.1016/S1473-3099(21)00249-8

[96] O. Westerland, A. Amlani, C. Kelly-Morland, M. Fraczek, K. Bailey, M. Gleeson, I. El-Najjar, M. Streetly, P. Bassett, G. J. R. Cook, V. Goh, London Myeloma Imaging Research Group at Guy's & St Thomas' Hospital, and King's College London. 2021. Comparison of the diagnostic performance and impact on management of 18F-FDG PET/CT and whole-body MRI in multiple myeloma. *European Journal of Nuclear Medicine and Molecular Imaging* 48 (2021), 2558–2565. doi:10.1007/s00259-020-05182-2

[97] Q. Xie, Y. Li, N. He, M. Ning, K. Ma, G. Wang, Y. Lian, and Y. Zheng. 2024. Unsupervised Domain Adaptation for Medical Image Segmentation by Disentanglement Learning and Self-Training. *IEEE Transactions on Medical Imaging* 43, 1 (2024), 4–14. doi:10.1109/TMI.2023.3324198

[98] X. Yang and J. Li. 2025. Learn from the SARS-CoV-2 nucleic acid test to increase the experience of dealing with the 'disease X'. *BMC Infectious Diseases* 25 (2025). https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-025-10991-7

[99] A. Dehghani Zahedani, T. McLaughlin, A. Veluvali, N. Aghaeepour, A. Hosseinian, S. Agarwal, J. Ruan, S. Tripathi, M. Woodward, N. Hashemi, and M. Snyder. 2023. Digital health application integrating wearable data and behavioral patterns improves metabolic health. *npj Digital Medicine* 6 (2023). doi:10.1038/s41746-023-00956-y

[100] Q. Zhang, Y. Wei, Z. Han, H. Fu, X. Peng, C. Deng, Q. Hu, C. Xu, J. Wen, D. Hu, and C. Zhang. 2024. Multimodal Fusion on Low-quality Data: A Comprehensive Survey. *arXiv preprint arXiv:2404.18947* (Nov. 2024). https://arxiv.org/abs/2404.18947

[101] Wenwu Zhu, Xin Wang, and Hongzhi Li. 2020. Multi-modal Deep Analysis for Multimedia. *IEEE Transactions on Circuits and Systems for Video Technology* (2020). doi:10.1109/TCSVT.2019.2940647

[102] X.-F. Zhu, T. Xu, Z. t. Liu, Z. Tang, X.-J. Wu, and J. Kittler. 2025. UniMod1K: A Dataset for Multimodal Learning with Vision, Depth, and Language. *International Journal of Computer Vision* 133, 6 (2025). https://link.springer.com/journal/11263/volumes-and-issues/133-6 Special Issue on Multimodal Learning.