Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Deep Generative Models Meet Federated Learning: A Healthcare-Centered Review

Alessandro Ceresi

alessandro.ceresi@upm.es

Universidad Politécnica de Madrid

Borja Arroyo-Galende

Universidad Politécnica de Madrid

Javier Guinea-Pérez

Universidad Politécnica de Madrid

Patricia A. Apellániz

Universidad Autónoma de Madrid

Gustavo Hernández-Peñaloza

Universidad Politécnica de Madrid

Federico Álvarez

Universidad Politécnica de Madrid

Research Article

Keywords:

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11484630/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Deep Generative Models Meet Federated Learning:

A Healthcare-Centered Review

Alessandro Ceresi[1*], Borja Arroyo-Galende[1], Javier Guinea-Pérez[1], Patricia A. Apellániz[2], Gustavo Hernández-Peñaloza[3], Federico Álvarez[1]

[1*]Information Processing and Telecommunications Center, Escuela Técnica Superior de Ingeniería de Telecomunicación, Universidad Politécnica de Madrid, Av. Complutense, 30, Madrid, 28040, Madrid, Spain.

[2]AUDIAS, Electronic and Communication Technology Department, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Av. Francisco Tomás y Valiente, 11, Madrid, 28049, Madrid, Spain.

[3]Escuela Técnica Superior de Ingeniería de Sistemas Informáticos, Universidad Politécnica de Madrid, Calle de Alan Turing, Madrid, 28031, Madrid, Spain.

*Corresponding author(s). E-mail(s): alessandro.ceresi@upm.es;
Contributing authors: borja.arroyog@upm.es;
javier.guinea.perez@upm.es; patricia.alonsod@uam.es;
gustavo.hernandez.penaloza@upm.es; federico.alvarez@upm.es;

Abstract

The integration of Artificial Intelligence (AI) in healthcare holds significant promise, yet remains constrained by data scarcity, privacy and ethical concerns, and the inherent complexity of medical tasks. This review explores the intersection of deep generative modeling and federated learning, two complementary paradigms that enable collaborative and privacy-preserving innovation. We survey a total of 85 articles on recent trends in deep generative models for healthcare, with a special focus on federated settings. The review is structured around model architectures, providing a taxonomy-driven analysis that emphasizes solutions tailored to the heterogeneity of healthcare data. The current literature

is largely limited to single-modality models; multimodal frameworks have only emerged over the last three years of our 7-year scope ($\boldsymbol{n = 14}$), where they are beginning to gain traction. Among deep generative approaches, generative adversarial networks dominate the field, particularly in medical image synthesis, where they comprise 66% of the studies. While alternative architectures such as variational autoencoders remain marginally represented in federated healthcare scenarios, diffusion and autoregressive models are gaining importance, suggesting a paradigm shift in the near future. Notably, the number of related publications has increased substantially in recent years, rising from 1 in 2020 to as many as 31 in 2025, with 2026 pointing toward a similar trend, reflecting the growing interest and momentum in this area. This paper highlights key challenges, open research questions, and future directions for the development of robust, distributed generative AI in healthcare.

Keywords: deep learning, federated learning, generative AI, healthcare, privacy

1 Introduction

Artificial Intelligence (AI) is revolutionizing healthcare by enabling faster and more accurate diagnoses, supporting informed clinical decision-making, and enhancing the optimization of healthcare delivery [1, 2]. Its capacity to process and learn from large-scale, complex medical data (such as Electronic Health Records (EHRs) [3], medical images [4], and signals from wearable devices [5]) is driving innovations across multiple domains, including early disease detection, patient monitoring, and personalized treatment planning. By supporting clinical workflows, AI contributes to more personalized, efficient, and effective care.

Despite these advances, the development and deployment of AI in healthcare remain hampered by critical challenges. These include the scarcity of labeled data, data imbalance, sampling bias, and the fragmented nature of medical records across institutions, affecting the quality of data [2]. Moreover, strict privacy regulations and ethical constraints further limit access to data, especially in cross-silo setups [6]. Such limitations are particularly acute in scenarios involving rare diseases, underrepresented patient subgroups, and low-resource clinical environments, where data are not only

scarce but also heterogeneous, non-standardized, and highly sensitive [7]. These issues undermine the generalizability of AI models, reduce reproducibility, and restrict their applicability in real-world clinical settings.

These challenges motivate the use of Deep Generative Models (DGMs), which aim to learn the underlying distribution of healthcare data and produce synthetic samples that preserve their structure and statistical properties [8]. Beyond alleviating data scarcity, DGMs are particularly well-suited to understand the complexity and heterogeneity of medical data, which often spans multiple modalities, variable formats, and intricate temporal dynamics. By capturing nonlinear dependencies and multimodal correlations, DGMs can generate realistic synthetic data that support data augmentation [9], fill in missing information [10], and enable model development when real data are limited or restricted [11]. In the context of EHRs, for example, DGMs have been used to synthesize patient trajectories that reflect clinically plausible event sequences while preserving privacy [12]. These capabilities make DGMs valuable not only for enhancing model robustness but also for facilitating research and system development in privacy-sensitive or resource-constrained environments.

However, while DGMs can help mitigate many data-related challenges, their training process still relies on access to real patient data, which raises critical concerns about privacy, data governance, and regulatory compliance (GDPR [13], EHDS [14], or HIPAA [15]). These concerns are particularly pressing in clinical contexts, where data cannot be freely shared across institutions or jurisdictions. In this context, Federated Learning (FL) has emerged as a powerful paradigm to enable collaborative model training across multiple entities without exchanging raw data [16]. Instead, models are trained locally on decentralized datasets, and only model updates —typically gradients or weights —are exchanged. By adopting a decentralized computational model, this approach boosts data privacy, ensuring that no single entity has full access to the complete dataset. Furthermore, FL enhances model robustness by exposing the

model to diverse data distributions across clients. Additionally, to mitigate the risk of information disclosure during training, FL can be extended with secure multi-party computation [17] or homomorphic encryption [18], which effectively limit the amount of information handled by the aggregator, even in the presence of semi-trusted or honest-but-curious participants.

The combination of DGMs and FL presents a compelling and increasingly relevant solution to the dual challenges of data scarcity and privacy in healthcare. While DGMs can alleviate the lack of accessible data by synthesizing realistic and downstream task-tailored samples, FL ensures that these models can be trained collaboratively across distributed institutions without exposing sensitive patient information. This integration is not merely complementary in function, but also in architecture: the parameters of DGMs, such as Generative Adversarial Networks (GANs) [19] or Variational Autoencoders (VAEs) [20], can be optimized in a federated fashion, preserving data locality throughout the learning process. Together, these paradigms reduce the need for centralized data sharing; however, empirical privacy assessments and supplementary defenses remain necessary to mitigate leakage during federated training and through generated outputs. Nevertheless, this synergy holds significant potential for unlocking collaborative research, building robust benchmarking datasets, and supporting the development of clinical AI in privacy-sensitive and data-constrained environments.

Inspired by these complementary strengths, this review explores their application to the healthcare domain. We focus specifically on deep neural architectures, which represent the current state of the art in terms of scalability and capacity to model high-dimensional, multimodal medical data. Traditional generative methods, while useful in constrained scenarios, lack the flexibility and expressive power required to capture the full complexity of clinical information or to generate coherent outputs across diverse data modalities [21, 22]. Our goal is to provide a comprehensive and

structured survey of recent works that apply DGMs in federated healthcare settings, highlighting architectural choices, methodological innovations, and application areas. In doing so, we identify common design patterns, key limitations in the existing literature, and open research questions that must be addressed to advance the development of robust, privacy-preserving generative AI in clinical practice. To facilitate a focused and comparative analysis, we adopt a data-centric perspective and organize the review according to the main data formats encountered in healthcare applications: tabular data, sequential data, and medical imaging.

The rest of this paper is structured as follows. In Section 2, we provide the necessary background on DGMs and FL, introducing key concepts and architectures relevant to their application in healthcare. Section 3 discusses the main challenges associated with different healthcare data modalities, highlighting how issues such as heterogeneity, privacy, and scarcity shape the design of generative solutions. Section 4 outlines the methods used to identify and select relevant studies for this review. Section 5 presents the results of the research, offering a structured overview of existing works along dimensions such as model type, data characteristics, and task formulation. Section 6 provides an analysis and discussion of the main findings of the reviewed literature, highlighting methodological patterns and research gaps. Finally, Section 7 concludes the paper by summarizing key insights and outlining open challenges and future research directions.

2 Background

This section lays the conceptual and technical groundwork for understanding the use of DGMs in healthcare, particularly when trained under privacy-preserving, distributed conditions. We begin by introducing the main families of DGMs and their relevance to medical data. We then discuss how these models can be trained collaboratively through FL to overcome data access constraints in clinical environments. Finally, we present key

dimensions and methodological challenges involved in evaluating the fidelity, privacy, and utility of generated data in healthcare applications. These foundations support the subsequent analysis of modality-specific challenges and recent developments in the field.

2.1 Foundations of Deep Generative Models

In data-driven modeling, DGMs aim to learn the underlying data-generating process of a given dataset [23]. Specifically, they estimate the probability distribution from which the observed data are assumed to originate [24]. Let x denote an observation drawn from an underlying data distribution, $x \sim p(x)$. A DGM parameterized by θ then learns an approximation of this distribution, denoted by $p_\theta(x) \approx p(x)$. More generally, DGMs can model joint or conditional distributions involving one or more variables. Let $x = (x^{(1)}, \ldots, x^{(M)})$ denote one or more data modalities and let c represent optional contextual or conditioning information. The generative model can then be expressed as $p_\theta(x, c) \approx p(x, c)$, or when conditioning information is available $p_\theta(x \mid c) \approx p(x \mid c)$.

DGMs can also be interpreted as graphical or hierarchical probabilistic models, where dependencies are defined between observed and latent variables [25]. These structures not only formalize the model architecture, but also provide a principled framework for sampling synthetic data [26]. Depending on the availability of data, DGMs can balance different degrees of prior assumptions. Models with strong prior structures are typically used in low-data regimes to inject domain knowledge, while models with weaker priors rely more heavily on learning from large datasets.

A wide range of DGM architectures has been developed in recent years. Within the scope of this review, we focus on four main families: GANs [19], Autoregressive (AR) models, VAEs [20], and Diffusion Models [27]. These approaches differ in their training objectives, generative mechanisms, and representational flexibility. A schematic summary of their training processes is provided in Figure 1. While other architectures,

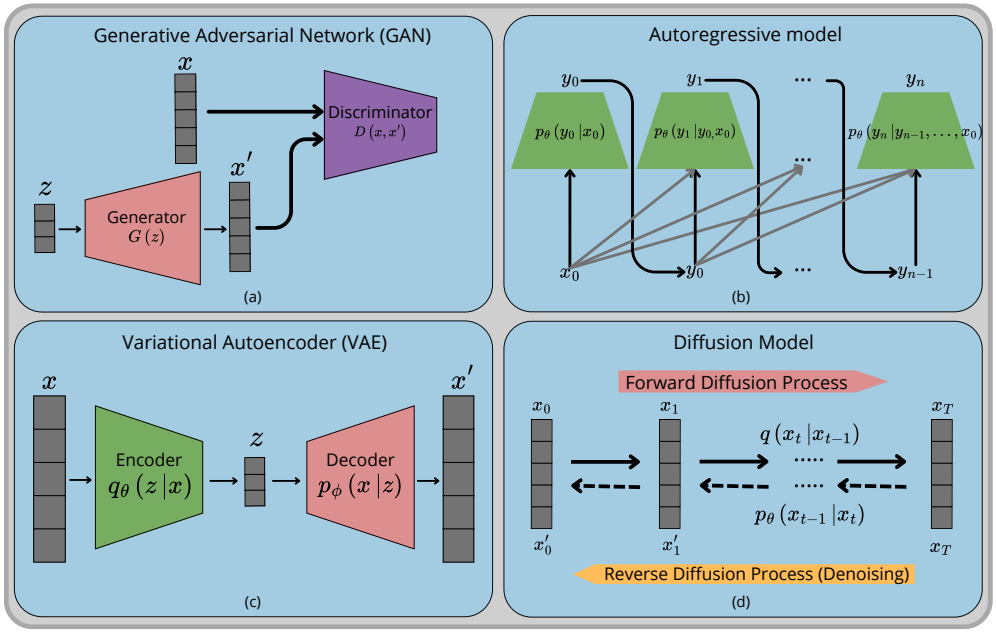


Fig. 1: Schematic representation of the core training processes of the four DGMs considered in this review: (a) **GANs**, where a generator produces synthetic samples that compete against a discriminator trained to distinguish real from fake data in an adversarial setup; (b) **AR models**, where data are generated sequentially (i.e., each output y_t is conditioned on previous outputs and inputs), and the model learns $p_\theta(y_t|y_{t-1}, ..., x_0)$, with θ as the parameter set; (c) **VAEs**, where the input data are encoded into a latent space and reconstructed through a decoder, q_θ approximates the posterior over latent variables, and p_ϕ reconstructs the input from the latent code z (θ and ϕ are the trainable parameters of the encoder and decoder); (d) **Diffusion models**, where generation involves a forward process that gradually adds noise to the data, and a reverse process learned to denoise it (q and p_θ denote the forward and reverse distributions, t is the timestep, and T the total number of steps).

such as normalizing flows [28] or energy-based models [29], are used in the literature for different types of healthcare applications, in FL, the resulting research is limited to these four architectures.

These models have shown significant potential in healthcare, particularly in addressing challenges such as data scarcity, heterogeneity, and privacy constraints [30, 31]. They enable the generation of realistic synthetic data that can be used for dataset anonymization, augmentation, domain translation, or imputation of missing

values [32, 33]. By preserving statistical and structural properties of the original data while protecting individual privacy [34], DGMs support the development of robust and generalizable AI models. They also facilitate clinical research in scenarios where real data are limited, restricted, or fragmented across institutions. Ultimately, DGMs can be tailored to simulate patient outcomes, support clinical decision-making, and personalize treatment planning [35], contributing to data-driven healthcare innovation while safeguarding patient confidentiality [36].

Generative Adversarial Networks (GANs)

GANs consist of two neural networks, a generator and a discriminator, that are trained in a competitive framework. The generator creates synthetic data, while the discriminator attempts to distinguish between real and generated samples [19]. Through this adversarial process, the generator learns to produce increasingly realistic outputs without explicitly estimating the underlying data distribution. This ability to synthesize realistic samples has made GANs particularly attractive in medical imaging and structured data generation. They are often used for data augmentation, class balancing, or generating synthetic cohorts when real patient data are limited or sensitive. However, their performance is often challenged by mode collapse (where the generator produces a narrow set of outputs), failing to capture the full diversity of the data distribution. This issue is exacerbated in healthcare settings, where datasets are often small, heterogeneous, and imbalanced.

Autoregressive (AR) Models

AR models generate data sequentially, where each element is predicted based on previously generated elements and, optionally, on additional inputs. Formally, they model the joint distribution of a sequence as a product of conditional probabilities. AR models are commonly implemented using recurrent neural networks (RNNs) or, more recently, transformer-based architectures. Large language models (LLMs), such as the

GPT family [37], are prominent examples of AR transformers. They have demonstrated remarkable capabilities in generating coherent, context-aware text, which has led to their adoption in tasks such as summarization, question answering, and dialogue generation. In healthcare, AR models are particularly relevant for modeling time-dependent data, such as patient trajectories, vital signs, or medication sequences, where preserving temporal consistency is essential [38]. Their ability to model sequential dependencies makes them well-suited for simulating longitudinal clinical data or forecasting future clinical events.

Variational Autoencoders (VAEs)

VAEs are probabilistic generative models that learn a mapping from input data to a lower-dimensional latent space and back, enabling the generation of new samples from a learned distribution [20]. They combine variational inference with autoencoding by optimizing a tractable lower bound on the marginal likelihood of the data. The encoder maps the data into a distribution over latent variables, while the decoder reconstructs the data from samples drawn from that distribution. VAEs are particularly valuable in healthcare for handling noisy, sparse, or incomplete datasets, where their latent representations can be used for imputing missing values, simulating patient trajectories, or learning compressed embeddings for downstream tasks. Their principled approach to handling uncertainty makes them attractive for privacy-sensitive applications, and they are among the most widely explored generative architectures in biomedical research.

Diffusion Models

Diffusion models are a class of DGMs that learn data distributions by progressively corrupting input data with noise and then training a neural network to reverse this process [27]. This approach provides a stable and flexible training mechanism that has recently yielded state-of-the-art performance in image and audio generation. The

forward process gradually adds noise over multiple timesteps, while the learned reverse process denoises the data to recover the original distribution. In healthcare, diffusion models are emerging as promising tools for generating high-fidelity synthetic data, particularly for complex modalities such as medical imaging and time series [39, 40]. They stand out in terms of stability and sample diversity, and their flexibility enables adaptation to a range of biomedical data types.

2.2 Federated Training of Deep Generative Models

Increasing concerns regarding the need for large, centralized datasets for training have accompanied the growing interest in generative modeling [41]. These concerns are especially critical in healthcare, where data are highly sensitive, regulated, and distributed across multiple institutions. Beyond the volume of data required, a key challenge lies in the secure, privacy-preserving use of this information across organizations that are unable (or unwilling) to share raw data due to legal and ethical constraints.

FL offers a compelling solution to this challenge by enabling collaborative training across decentralized nodes without transferring raw data [42]. In a typical FL setup, each data provider (e.g., a hospital or data silo) trains a local model using its private data. These local models are periodically merged by a central aggregator, which collects and combines model updates, such as parameters or gradients, using algorithms such as Federated Averaging (FedAvg) [43]. This paradigm preserves data locality while enabling the training of a shared global model.

Let $\mathcal{K} = \{1, \ldots, K\}$ denote a set of K clients, where client k holds a private dataset

$$\mathcal{D}_k = \{(x_{k,i}, c_{k,i})\}_{i=1}^{n_k},$$

with $x_{k,i}$ denoting the observed data and $c_{k,i}$ denoting optional conditioning information. More generally, let $x = (x^{(1)}, \ldots, x^{(M)})$ represent a collection of M modalities.

A federated generative model aims to learn a global distribution $p_\theta(x \mid c)$, or equivalently, a joint distribution $p_\theta(x, c)$ depending on whether conditioning information is explicitly provided at generation time. This formulation encompasses unconditional generative models by setting $c = \varnothing$, conditional generative models through $p_\theta(x \mid c)$, and multimodal generative models by defining x over multiple modalities.

In the federated setting, the global model is learned by optimizing local objectives,

$$\min_\theta \ \mathcal{L}(\theta) = \sum_{k=1}^{K} w_k \, \mathcal{L}_k(\theta; \mathcal{D}_k), \qquad w_k = \frac{n_k}{\sum_{j=1}^{K} n_j},$$

while the underlying client datasets remain private. After training, the global model can generate samples according to

$$\hat{x} \sim p_\theta(x \mid c),$$

for a given condition c, or from $p_\theta(x)$ in the unconditional case.

This paradigm is particularly valuable in privacy-sensitive domains such as healthcare, where data often include confidential patient information [44]. FL provides a secure way to comply with strict data protection regulations, like GDPR 2016/679 [45] or HIPAA [46]. To further strengthen privacy guarantees, FL can be combined with techniques such as differential privacy (DP), which adds noise to sensitive data to obscure individual contributions [47], and secure multiparty computation, which allows encrypted collaboration between nodes [48].

However, applying FL to DGMs introduces unique technical challenges not present in discriminative tasks. While predictive models learn a conditional distribution $p(y|x)$, generative models aim to learn the joint distribution $p(x)$ (or related structured distributions), making them more sensitive to inconsistencies across nodes. In particular, FL for DGMs must contend with non-independent and identically distributed (non-IID) data, which is the norm in healthcare, as hospitals often treat different patient populations, use heterogeneous equipment, and follow varying clinical protocols. This

heterogeneity can degrade both model convergence and the fidelity of generated samples [49]. Moreover, DGMs often require rich latent structures that generalize well across institutions. Maintaining consistency in the learned latent space under non-IID conditions is particularly difficult, and naive aggregation may lead to mode collapse, over-smoothing, or loss of diversity in generated outputs. These limitations are further amplified in federated settings by limited communication bandwidth, variable local computational resources, and the need for synchronization between nodes.

Despite these challenges, combining FL and DGMs has clear potential for healthcare: it enables privacy-preserving data synthesis, augmentation of underrepresented classes, generation of realistic rare cases, and simulation of patient trajectories, all without compromising patient confidentiality. This makes federated generative modeling a promising but still nascent research area in the development of trustworthy AI systems for clinical use [50].

2.3 Evaluation Dimensions for Deep Generative Models in Healthcare

Evaluating DGMs remains one of the most challenging and unsettled aspects of their development, particularly in sensitive domains such as healthcare [51]. Unlike discriminative models, for instance, which are typically assessed by their predictive performance on a test set (e.g., by estimating $p(Y|X)$ and computing classification or regression metrics), DGMs aim to learn the underlying joint distribution $p(X)$ and generate new samples from it. This difference makes direct evaluation more complex, especially when the output is not a single prediction but an entire synthetic dataset.

Currently, there is no universally accepted standard for evaluating DGMs [52, 53]. Instead, evaluation strategies are typically designed based on the intended use case, such as imputation, data augmentation, or anonymization, and must account for both technical accuracy and ethical implications [54–56]. In healthcare, where data are often

scarce, heterogeneous, and sensitive, this evaluation becomes even more critical. To address this, most frameworks adopt a multi-dimensional evaluation approach based on three core aspects: fidelity, privacy, and utility [57, 58].

- **Fidelity** refers to the degree to which generated data replicate the statistical properties, structure, and complexity of the original data distribution. It ensures that the generated samples are realistic and coherent [59].
- **Privacy** assesses the risk of re-identification or information leakage from generated data, ensuring that these samples do not expose sensitive details about individuals in the original dataset [59]. In medical settings, this is crucial to ensure regulatory compliance and ethical integrity.
- **Utility** evaluates whether generated data can serve as a meaningful substitute for real data in downstream tasks, such as model training or exploratory analysis [60]. It reflects the practical value of these data in real-world clinical or research applications.

A balanced assessment across these three dimensions is essential for judging the overall quality and reliability of DGMs in healthcare. However, trade-offs often arise (for example, improving fidelity may compromise privacy [61]), making evaluation not only a technical challenge but also a matter of design choice aligned with specific clinical or regulatory priorities.

2.4 Privacy, Security, and Regulatory Considerations

Beyond technical design, deploying DGMs in healthcare also requires careful attention to privacy, security, and regulatory compliance. As these models increasingly approach clinical-grade applications, particularly in areas such as medical imaging or clinical decision support, understanding the legal and ethical landscape becomes essential. This subsection outlines the regulatory context and associated security challenges, along with key mitigation strategies.

Recent policy initiatives, most notably the FDA's executive summary *"Total Product Lifecycle Considerations (TPLC) for Generative AI-Enabled Devices"* [62] and the European Union's AI Act [63], underscore the necessity of holistic governance for DGMs under the domain of medical devices. These frameworks emphasize lifecycle-oriented governance, calling for rigorous design validation, quality assurance, and continuous post-deployment monitoring. In particular, the FDA highlights the unique challenges posed by generative AI in clinical applications, including variability of outputs and model drift, and emphasizes the need for robust oversight mechanisms.

Among the regulatory challenges identified, three stand out in particular. First, the variability of generative outputs complicates validation processes, since models may produce a wide range of plausible results rather than a single deterministic outcome. Then, the rapid pace of model iteration demands continuous oversight and evaluation, making static approval pipelines insufficient for ensuring long-term safety and reliability. Moreover, third, the data quality and bias remain critical, as biased or low-quality datasets can propagate errors and inequities into clinical workflows.

Privacy-enhancing technologies such as FL and DP are specifically endorsed by policy frameworks, including the EU AI Act, as proactive compliance mechanisms to mitigate privacy and security risks. FL enables decentralized collaborative training without direct data sharing, aligning with data minimization principles mandated by regulations like GDPR and HIPAA.

However, these same techniques introduce new security vulnerabilities that must be proactively managed. DGMs risk inadvertently memorizing sensitive training data, making them susceptible to membership inference attacks [61, 64]. DP, although recommended, poses a privacy-utility trade-off [65]. In addition, FL, while mitigating direct data sharing, is not immune to adversarial behavior. Malicious clients may introduce compromised model updates (poisoning) or reconstruct sensitive data from aggregated gradients (inversion attacks) [66, 67]. Another important concern is

model stealing: DGMs represent valuable intellectual property, vulnerable to theft through unauthorized querying and replication. Current countermeasures, such as watermarking, are promising but still evolving [68].

Effectively mitigating these security risks while leveraging FL and privacy-preserving techniques ensures compliance with comprehensive lifecycle-based regulatory frameworks. Robust defense mechanisms, continuous threat monitoring, and adherence to international standards (such as ISO/IEC 27001, ISO/IEC 42001, and the NIST AI Risk Management Framework) become essential for successful regulatory alignment.

3 Challenges Across Healthcare Data Modalities

Traditionally, data have been treated as homogeneous entities. However, in practice, healthcare information exists in various modalities, each characterized by distinct structural and statistical properties. In this context, a data modality refers not only to the format in which data are represented, but also to the internal coherence and dependencies that define its semantics. For example, images encode spatially correlated information [69], while sequential data, such as clinical notes or physiological signals, unfold over time and require temporal modeling [70].

The goal of this section is to examine how these modality-specific characteristics impact the development and deployment of DGMs in healthcare. We structure the discussion around four major data categories: tabular, sequential, imaging, and multimodal. For each, we outline its clinical relevance, describe the structural and operational constraints that complicate generative modeling, and highlight the unique challenges that arise when such data are handled in privacy-sensitive, decentralized environments. This analysis provides the foundation for understanding both the opportunities and limitations of current approaches and for identifying modality-specific gaps in the literature.

3.1 Image data

Medical imaging plays a crucial role in clinical workflows, enabling the visualization of internal anatomical structures for diagnosis, monitoring, and treatment planning. Techniques such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and digital pathology produce high-resolution data essential for understanding complex physiological conditions. In this context, the generation of high-quality synthetic medical images has emerged as a key area of research, offering a solution to data scarcity, class imbalance, and privacy concerns in model training [71].

However, working with medical image data in federated generative frameworks introduces several significant challenges. The storage requirements of inherently high-dimensional and large imaging data pose a challenge for local training and communication in federated frameworks, especially for 3D volumes or whole-slide images. Variability across scanners, acquisition protocols, and institutions further complicates model generalization, while preserving anatomical coherence and pathology-specific details remains a challenge. Limited and imbalanced annotations, due to privacy restrictions, costs, and disease rarity, also restrict data representativeness in federated scenarios [72]. Furthermore, gradients from high-resolution images can leak sensitive information if not adequately protected, and training large-scale DGMs, such as diffusion models or GANs, often demands significant GPU resources. In federated environments, this can exceed the capabilities of some clinical nodes or require special scheduling and orchestration mechanisms.

3.2 Tabular data

A significant portion of the studies considered in this review relies on tabular data, which remains the dominant format for representing structured clinical information. This includes patient demographics, laboratory results, vital signs, diagnostic codes,

treatment history, and genomic variants. Tabular data power a wide range of tasks, from risk prediction to clinical decision support and health economics modeling. Generating synthetic tabular data that preserve the complex statistical and clinical relationships present in real datasets is crucial for enabling privacy-preserving research and robust model development [73–75].

Despite their structured format, tabular healthcare data present distinct challenges for generative modeling, especially in federated settings. These datasets often contain hundreds or even thousands of features per patient. Yet, many variables are sparsely populated due to irregular data collection, missing values, or the rarity of certain conditions [76]. Such sparsity makes it difficult to estimate realistic joint distributions, particularly in low-data or non-IID scenarios. Moreover, tabular data include numerical, ordinal, categorical, and binary variables, often with multimodal or skewed distributions. Building generative models that can simultaneously capture these discrete–continuous mixtures, preserve inter-variable dependencies, and respect logical constraints (e.g., ensuring that medical codes are compatible with age and sex) remains an open problem. These difficulties often result in incomplete mode coverage, particularly for underrepresented subgroups or rare feature combinations. Assessing the plausibility and coherence of generated records introduces another layer of complexity, since clinical expertise is often required for validation. Finally, in federated environments, heterogeneity across institutions in EHR systems, data codification, variable naming, and measurement units complicates the alignment of features and model consistency. Without careful harmonization, models risk capturing site-specific artifacts, producing biased or non-generalizable outputs. Moreover, when generating synthetic data for rare disease cohorts or small subpopulations, the risk of re-identification increases if models overfit or memorize individual records. While federated frameworks can reduce this risk, they must be complemented with additional safeguards to ensure privacy.

3.3 Sequential Data

Several healthcare data modalities are inherently sequential, capturing processes that evolve. These include biomedical time-series (e.g., electrocardiogram (ECG), electroencephalogram (EEG), or accelerometer signals), temporally indexed clinical markers (e.g., red blood cell count, medication history), unstructured text sequences (clinical narratives or event logs), and biological sequences (e.g., DNA or RNA). These data types are characterized by temporal or positional dependencies, variable sequence lengths, and often exhibit irregular sampling or censoring due to practical constraints in data collection.

Generative modeling of sequential data, especially in federated settings, introduces a range of technical challenges. Clinical sequences, such as laboratory measurements or symptom progressions, are frequently sampled at non-uniform intervals, which breaks the assumptions of standard AR or RNN-based architectures and requires models that explicitly handle time gaps or irregular updates. Longitudinal datasets also suffer from censoring and incomplete records, especially in survival analysis and EHR logs, making it essential for DGMs to simulate realistic patient trajectories under partial observability and uncertain event timing. Furthermore, sequential data are highly heterogeneous: biosignals are continuous and high-frequency, while genomic sequences are discrete and symbolically ordered. Designing models that generalize across such variable-length and structurally diverse sequences remains an open challenge. Another key difficulty lies in generating plausible and clinically coherent trajectories. Synthetic ECG waveforms, for example, must preserve valid morphologies, while simulated treatment sequences should reflect realistic therapeutic pathways. In federated environments, additional complications arise due to temporal drift and non-IID distributions, as data collected across institutions or time periods may reflect shifts in patient populations or clinical practices. This heterogeneity can degrade model generalization and destabilize the

training process. Finally, sequential data tend to be large due to high-frequency sampling or long sequence lengths, increasing communication costs in FL. Although latent compression techniques such as VAEs can reduce bandwidth requirements, they risk discarding clinically relevant temporal details [77].

3.4 Multimodal Integration

In real-world healthcare scenarios, information about a patient is rarely captured through a single type of data. Instead, multiple sources (each with different structures, semantics, and temporal resolutions) are often combined to build a more complete representation of clinical states. This is the case, for instance, with EHRs, which integrate structured tabular data (e.g., demographics, lab results), unstructured text (e.g., clinical notes), time-series signals (e.g., heart rate), and medical images [78, 79]. These heterogeneous, longitudinal records present a rich yet challenging landscape for AI models, particularly for DGMs.

Each modality presents its specific challenges (refer to previous sections), requiring specialized architectures tailored to their unique statistical properties [80]. Beyond the complexity within each modality, combining them introduces a new layer of difficulty: multimodal models must learn both intra-modality and cross-modality relationships, increasing architectural complexity and optimization difficulty [81–83]. In healthcare, this fusion holds the promise of revealing richer clinical patterns and improving prediction performance [84]. For instance, combining genomic features with imaging or biosignal data can uncover causal relationships that may inform disease progression or treatment response [85].

Multimodal generative modeling aims to synthesize or complete datasets that span two or more modalities. This can involve tasks such as domain translation (e.g., generating pathology images from genomic profiles), joint generation (e.g., producing

aligned images and reports), or unified synthesis from latent representations. In healthcare, these models are particularly appealing because they enable the simulation of missing modalities, augmentation of rare multimodal cases, and causal exploration across biological systems.

Nonetheless, developing multimodal DGMs, particularly under federated constraints, involves a series of challenges. A primary difficulty is the balance of modality and partial availability, as not all institutions collect the same types of data. Some patients may lack imaging, while others have no molecular data. This imbalance complicates training and reduces downstream utility, especially in federated setups where client data are inherently incomplete. Moreover, multimodal data often suffer from asynchrony and resolution mismatches: modalities differ in spatial, temporal, or semantic resolution, making it difficult to align inputs for generation. Examples include combining high-resolution imaging with sparse clinical codes or continuous vital signs. Another major challenge is achieving joint representations and semantic coherence; multimodal DGMs require shared latent spaces that capture intra- and inter-modality dependencies while ensuring that generated modalities remain semantically consistent (e.g., an image matching its synthetic clinical report) [83, 86]. Equally important is the need for conditional and flexible generation, where models must be capable of synthesizing one modality conditioned on others, even when combinations vary across patients or federated sites [81]. Heterogeneous client capabilities further complicate this, since institutions differ in both available data and computational resources. Training large-scale multimodal models locally may be infeasible at some nodes, requiring coordinated orchestration or fallback strategies. Finally, multimodal integration raises cross-modal privacy risks: while individual modalities may appear anonymized, their combination can amplify re-identification risks. Multimodal DGMs must address this compounded privacy concern in both centralized and federated deployments.

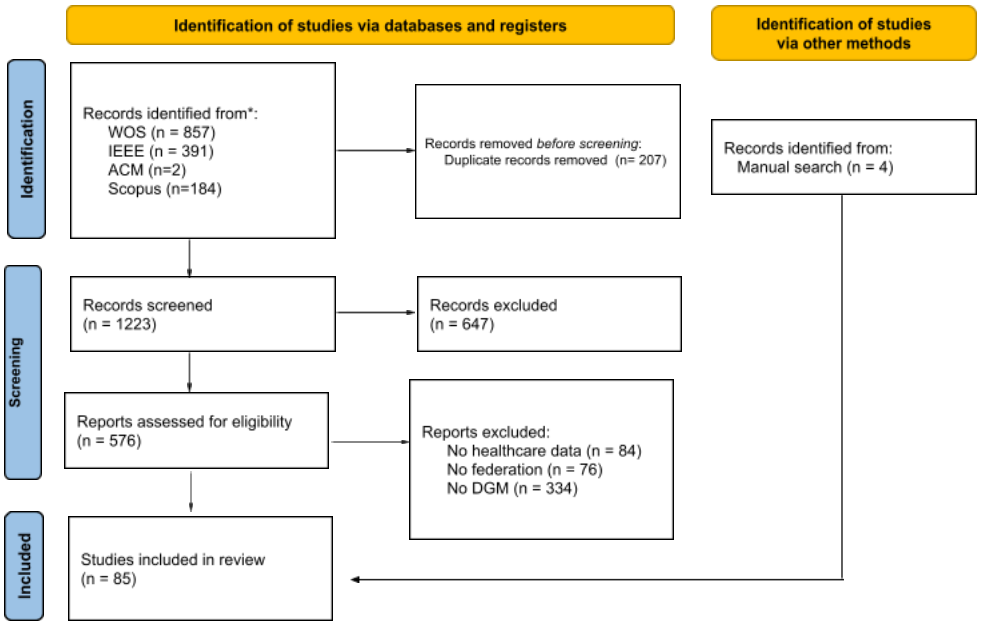


Fig. 2: PRISMA procedure workflow.

In summary, multimodal data offer the richest view of patient health but also represent one of the most complex modeling challenges for generative AI. Their inherent heterogeneity, incomplete structure, and high dimensionality demand architectures that are both flexible and privacy-aware, especially when used in distributed or cross-institutional settings such as FL.

4 Search Methodology

This section describes the methodology employed to identify, select, and analyze the scientific literature reviewed in this work. The process included the systematic collection of publications, keyword-based querying, and multi-stage filtering to ensure relevance to the use of DGM techniques in federated healthcare contexts.

Table 1: Summary of the query design according to the keywords

Engine	Results	Query	Comments
WOS	857	TS=((Federated OR Federative) AND (Generativ* OR Synthesiz* OR Variational OR Adversarial OR GPT* OR GAN OR Generation OR VAE OR Autoregressive OR Diffusion) AND (Health* OR Medic*))	Also limited the publication year and document type to remove reviews.
IEEE	391	((((((”Abstract”:Generat*) OR (”Abstract”:Synthesiz*) OR (”Abstract”:Variational) OR (”Abstract”:Adversarial) OR (”Abstract”:Generation) OR (”Abstract”:GPT*) OR (”Abstract”:GAN) OR (”Abstract”:Autoregressive) OR (”Abstract”:VAE) OR (”Abstract”:Diffusion)) AND ((”Abstract”:Health*) OR (”Abstract”:Medic*)) AND (”Abstract”:Federat*)))))	Also limited publication year and document type to journal entry.
ACM	2	[Abstract:federat*] AND ([Abstract:health*] OR [Abstract:medic*]) AND [Keywords:generat*] AND [E-Publication Date:(01/01/2020 TO 03/31/2026)]	Did not include particular architectures as it added too much noise.
Scopus	184	(((”Index Terms”:generat*) AND ((”Abstract”:health*) OR (”Abstract”:medic*)) AND (”Abstract”:federat*))) AND PUBYEAR>2020 AND PUBYEAR<2026	Did not include particular architectures as it added too much noise.

An exhaustive investigation was conducted on several relevant search engines, including Web of Science[1], IEEE Xplore[2], Scopus[3], and ACM Library[4]. Queries were built using carefully selected keywords to ensure the exploration of all existing literature, allowing the identification of pertinent works for this review.

In addition to the previously indexed sources, we also manually included a small number of relevant *preprints* hosted on arXiv. Since arXiv[5] is not covered by the

[1] Web of Science: https://www.webofscience.com. Accessed April 9th, 2026.
[2] IEEE Xplore: https://ieeexplore.ieee.org/. Accessed April 9th, 2026.
[3] Scopus: https://www.scopus.com/. Accessed April 9th, 2026.
[4] ACM Digital Library: https://dl.acm.org. Accessed April 9th, 2026.
[5] ArXiv: https://arxiv.org/. Accessed April 9th, 2026.

above digital libraries' indexing, these papers were added either because the authors were already aware of them or because they surfaced through supplementary, ad-hoc keyword searches on arXiv directly.

The overall process was carried out in a three-phase approach as presented in Figure 2, following the guidelines established by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The search words cover the main topics, and the search was aligned with a set of filters defined to aid in the identification of relevant works. The procedure is described in the following subsections.

4.1 Identification: General Searching

Initially, general terms related to the use of DGMs were considered to define the scope of the study. Consequently, the terms generative, generation, etc., were included, also synthesize-rooted terms. To broaden the search and include papers in which a particular architecture was presented but none of the generation or synthesis terms appear, the most common architectures were included in the search. To limit the scope of the search to the healthcare domain, either medical or healthcare-related terms had to occur in the papers. Another search requisite was the appearance of terms related to FL. Thus, papers meeting all three search criteria: FL, data generation, and the healthcare domain, were listed. To limit our search to the state of the art, only papers that were published between 2020 and March 31, 2026 were included.

To unify the search criteria, a total of 4 similar queries were designed for the selected digital libraries. Each query was formulated to intersect three dimensions: research focus, specific healthcare applications, and methodology for AI-based data generation. Table 1 provides a summary of the query results, outlining the scope and outcomes of the search. The queries were formulated to be as similar as possible, given the limitations of the query formulation tools available.

Because of the overlapping contents of the databases queried, duplicate records were removed before beginning the screening process.

4.2 Screening

Once the entire list was obtained, the process focused on identifying and discarding unrelated and out-of-scope papers. The objective was to refine the list and narrow it down to the most relevant papers for appraisal. Some unrelated papers complied with the search criteria, and other related papers were either not employing FL or not generating data per se. To discard these papers, the following criteria were applied:

- **Screening process**: To narrow the number of papers included in the eligibility assessment, a first screening was performed by individual examination of the title and abstract of each article. Papers that met the search criteria but were deemed irrelevant to the topic were subsequently excluded from the review. For example, articles discussing next-generation mobile networks or articles where healthcare or FL were only mentioned in the motivation or as an example.
- **Eligibility criteria 1, Healthcare Data**: At this stage, a general inspection of each paper was conducted, and articles were excluded if their experimental setup did not include healthcare data.
- **Eligibility criteria 2, Federation**: In this step, papers were filtered based on the federation strategy they followed. Papers in which no federation was performed were excluded from the review.
- **Eligibility criteria 3, DGMs**: Lastly, papers in which generation was stated but no generative model or task was present were removed. This included papers that used non-generative models, such as autoencoders.

The following section presents the results of the overall filtering process.

Table 2: Overview of federated applications of DGMs in Healthcare.
The generative tasks categories were defined as follows: *Synthesis* as default category, *Augmentation* when synthetic data are used to enrich a scarce training dataset, *Balancing* when synthetic data are generated to balance underrepresented classes of a dataset, *Domain Translation* when synthetic data are generated in a domain different from that of the real data, *Imputation* when the generation process is employed to fill in missing or incomplete parts of a real dataset. The color code represents the replicability of the article methods: both code AND data are available, either code OR data are unavailable, both are unavailable.

Architecture	Data type	Generation task	Non IID	Reference
AR	Image	Synthesis	✓	[87][88][89][90][91]
			✗	[92]
	Image, Sequential	Domain Translation	✓	[93]
		Synthesis	✗	[94][95]
	Image, Tabular	Synthesis	✓	[96]
	Sequential	Synthesis	✓	[97][98][99]
			✗	[100]
	Sequential, Tabular	Synthesis	✓	[101]
Diff	Image	Augmentation	✓	[102][103]
		Domain Translation	✓	[104]
		Synthesis	✓	[105][106][107][108][109]
			✗	[110][111][112]
	Tabular	Synthesis	✓	[113]
Diff + AR	Sequential, Tabular	Synthesis	✓	[114]
GAN	Image	Augmentation	✓	[115][116][117][118]
			✗	[119][120][121][122][123][124] [125][126][127][128][129]
		Balancing	✓	[130][131][132]
		Domain Translation	✓	[133][134][135][136][137][138] [139][140][141][142]
			✗	[143][144]
		Synthesis	✓	[145][146][147][148]
			✗	[149][150][151]
	Image, Tabular	Augmentation	✓	[152]
		Imputation	✓	[153]
	Sequential	Augmentation	✗	[154]
		Domain Translation	✓	[155]
	Sequential, Tabular	Imputation	✗	[156]
	Tabular	Augmentation	✗	[157]
		Synthesis	✓	[158][159]
			✗	[160][161]
GAN + AR	Image, Sequential, Tabular	Augmentation	✗	[162]
	Sequential	Domain Translation	✗	[163]
GAN + Diff	Image	Domain Translation	✓	[164]
GAN, VAE[1]	Image, Sequential, Tabular	Synthesis	✓	[165]
VAE	Image	Augmentation	✓	[166][167]
	Sequential	Augmentation	✗	[168]
		Balancing	✓	[77]
			✗	[169]
	Tabular	Augmentation	✓	[170]

[1] The articles in this category employ different model architectures in separate experiments, instead of combining them.

5 Results

The results of this review are organized based on the architecture of the DGMs proposed in the literature, reflecting the highly specialized nature of these models within FL environments for healthcare applications. This approach is motivated by the observation that the applications of DGMs in healthcare are remarkably diverse, spanning tasks such as synthetic data generation, data augmentation, dataset balancing, domain translation, and privacy-preserving data sharing. Each use case often demands tailored solutions, with the architecture of the DGM being central to addressing the specific requirements and challenges of the task. Given this strong dependence on architectural design, categorizing the reviewed studies by their DGM type, such as GANs, VAEs, diffusion models, or AR models, provides a clearer framework for understanding their methodologies, innovations, and limitations. This structure also facilitates comparisons across studies employing similar architectures, highlighting trends and gaps in the field, which will be discussed in the following sections. All articles selected for this review are categorized in Table 2; furthermore, they can be filtered, grouped, and examined in greater detail via the accompanying interactive table[6].

5.1 Generative Adversarial Networks (GANs)

In the landscape of federated generative modeling for healthcare, GANs emerge as the most prevalent architecture for mitigating data scarcity, class imbalance, and multi-institutional privacy barriers. Technical implementations typically evolve from vanilla GAN architectures coordinated by FedAvg toward more stable local generative frameworks like Wasserstein GANs with Gradient Penalty (WGAN-GP) to counter mode collapse and training instability inherent in medical datasets [119, 125, 129, 159, 161]. Additionally, conditional GANs (CGANs) provide targeted synthesis of underrepresented pathologies by incorporating auxiliary information into the generation

[6] An interactive and extended version of the table is available at: https://gatv.gitlab.io/synthia/federated-gm-in-healthcare-review/

[130, 137, 145, 158]; while CycleGANs enable unpaired cross-modality translation [119, 144, 151]. Another architectural trend is to decouple generator and discriminator, keeping discriminators local to prevent patient-specific leakage due to gradient updates.

High-dimensional imaging remains the primary modality, with frameworks specializing in MRI diagnosis or reconstruction such as FAME and FedGIMP [126, 127, 130, 134, 143], while others focus on chest X-ray and CT synthesis for COVID-19 and pneumonia detection [117–122, 131, 132, 146, 147]. Advanced implementations further extend to intermodality domain translation and image-to-image mapping, such as MRI-to-CT transitions for clinical planning or the synchronization of datasets across varying hardware protocols and modalities [133, 134, 136, 140–144, 164]. Specialized tasks have also expanded to include the 3D and 4D segmentation of complex structures, ranging from brain tumours and respiratory lesions to cardiovascular anatomical components, ensuring diagnostic consistency across heterogeneous institutions [128, 129, 133, 138–141, 150, 151]. Federated GANs are also extensively applied to texture-centric 2D modalities, where synthetic augmentation serves as the principal technical tool for mitigating severe class imbalances and data scarcity in specialized skin disease datasets [116, 117, 123–125, 135, 145, 148, 149, 153]. Beyond skin lesion synthesis, these frameworks are increasingly adapted for a range of other texture-centric tasks to enhance diagnostic precision in other fields, such as ophthalmology and neurology [115, 117, 137–139, 146]. To address non-visual data modalities, federated GANs have been adapted to generate high-fidelity tabular EHRs. Specialized frameworks such as DP-CTGAN and FedCG-WGAN prioritize the generation of discrete clinical indicators and statistical patient records, for tasks ranging from cancer detection and gout staging to Alzheimer's disease classification [157–161]. Furthermore, federated GANs are increasingly adapted for processing longitudinal biometric

signals, such as continuous blood pressure estimation [163], lung sound classification [154], and multi-task monitoring of physiological data like EEG and ECG [155]. Finally, emerging multimodal frameworks provide a holistic diagnostic view by integrating these heterogeneous sources, such as combining clinical notes with structured vitals for sepsis detection or fusing patient metadata with medical images through attention-guided co-training [152, 156, 162, 165].

5.2 Autoregressive (AR) models

One of the main challenges in federating AR models is data throughput. AR models can have a large number of parameters, making updating all parameters unfeasible in federated environments. The most common solution is to employ a static model and use lightweight adapters between some layers that are then federated [116], using Low-Rank adaptation (LoRA). Another complementary approach is to utilize prompt and prefix-tuning [90], where a trainable prompt embedding is injected into the model's input or attention layers, leaving the backbone frozen. Some papers share high-level feature summaries (prototypes) instead of the whole model or adapters [88, 93] to further reduce the number of parameters shared. To reduce communication costs, some authors quantize models [94, 97], share only some of the gradients [94] or share only the differentials of the parameters changed on each checkpoint [100]. Lastly, researchers can opt to use more compact models [100] or use pretrained foundational models (such as Med-PaLM [171] or BioGPT [172]) to reduce the training effort [95, 100].

Almost all AR models in the literature included in this review are aimed at text generation. Two are focused only on using LLMs to share medical knowledge through a chatbot [100] and for clinical decision support [97]. Contemporary research has moved toward multimodal foundational models, using vision transformers as encoders and LLMs as decoders, with the AR decoders generating human-readable outputs of visual tasks, such as visual question answering [90, 93, 94, 99] and report generation or

interpretable diagnosis [87–89, 91, 92, 95, 96, 99]. Only one paper in the review does not generate text as output: Kim et al. [101] use a Bidirectional Encoder Representations from Transformers (BERT) inspired architecture for hypertension management, where future systolic blood pressure values are generated using old values and clinical data.

5.3 Variational Autoencoders (VAEs)

Despite the widespread application of VAEs in centralized settings, their presence in FL environments, where data privacy, distributional heterogeneity, and communication efficiency are critical, is limited.

Among the reviewed literature, only seven healthcare studies stand out for explicitly focusing on a VAE architecture within a federated framework, making it the least represented generative architecture in this domain. In the reviewed works, researchers primarily employ the VAE's generative capacity to mitigate data imbalance and distributional shifts. Zhao et al. [166] introduced the FedDIS approach for medical image classification, which utilizes a VAE to estimate and share hidden space distributions, allowing clients to augment local data and effectively reach IID conditions. Similarly addressing heterogeneity, Apellániz et al. [170] integrated Bayesian Gaussian Mixture models into the VAE latent space to capture complex tabular data distributions, demonstrating that sharing synthetic patient records can significantly reduce performance disparities between diverse healthcare institutions. This focus on distribution balancing is also evident in the work of Sikandar et al. [169], who used VAE-based generative oversampling to resolve extreme class imbalances in genomic sequences for cancer detection. Beyond data augmentation, the VAE architecture is increasingly adapted to enhance communication efficiency and deep personalization. Yang et al. [77] proposed the FedVAE framework for remote patient monitoring, leveraging lightweight VAEs to project high-dimensional data into low-dimensional latent features. In specialized biomechanics research, Huang et al. [165] utilized VAEs within a federated

framework to analyze cellular deformation patterns, enabling multi-institutional collaboration while mitigating the re-identification risks of sensitive biomechanical data. To resolve more complex domain and label shifts in image analysis, Lakrouni et al. [167] introduced the GPAF framework, which learns domain-invariant representations by using a global generator to guide the adversarial alignment of local conditional distributions within the latent space. Finally, to bridge the gap between collective knowledge and individual variability, Lv et al. [168] introduced Med-FedLSG, which employs LSTM-based variational mechanisms and Federated PCA to explicitly disentangle the latent space into shared and personalized subspaces for wearable health monitoring.

5.4 Diffusion models

While the integration of diffusion models into FL frameworks was previously limited to a few experiments, they have recently emerged as an imposing standard in healthcare, with a rapidly increasing number of works published every year.

Several studies prioritize client-side generative augmentation using Latent Diffusion Models (LDMs) [173] or Denoising Diffusion Probabilistic Models (DDPMs) [27] to rectify statistical heterogeneity in classification tasks, as demonstrated in the works of Li et al. [102], Huangsuwan et al. [105], Zhou et al. [103], and Tun et al. [110]. Architectural innovations have been specifically tailored for structural and segmentation precision, including mask-guided cascaded models for multi-center scenarios (Zhang et al. [107]), highlight diffusion mechanisms designed to accentuate blurred anatomical boundaries (Niu et al. [104]), and attention-guided modules prioritizing discriminative features for Alzheimer's diagnosis (Song et al. [109]). For specialized medical modalities, researchers have incorporated Transformer-driven generators for molecular discovery (Hou et al. [114]), U-ViT backbone diffusion for distilled histopathology image sets (Li et al. [111]), and architectures designed for mixed-type tabular record

synthesis (Sattarov et al. [113]). Beyond standard fidelity, advanced conditioning techniques like client-specific vector quantization (Yoon et al. [108]) and weakly supervised learning for federated continual diagnosis (Sun et al. [106]) have been introduced to address local distribution drift. Furthermore, recent frameworks have integrated specialized adaptation protocols like DreamBooth for robust pathological feature binding (Li et al. [102]) and one-shot aggregation methods to minimize communication overhead (Zhang et al. [107]). Finally, diffusion architectures are utilized both for eliminating voxel-wise distribution shifts between imaging scanners (Li et al. [164]) and for evaluating system vulnerabilities through deep generative gradient inversion attacks (Wu et al. [112]).

6 Analysis and Discussion

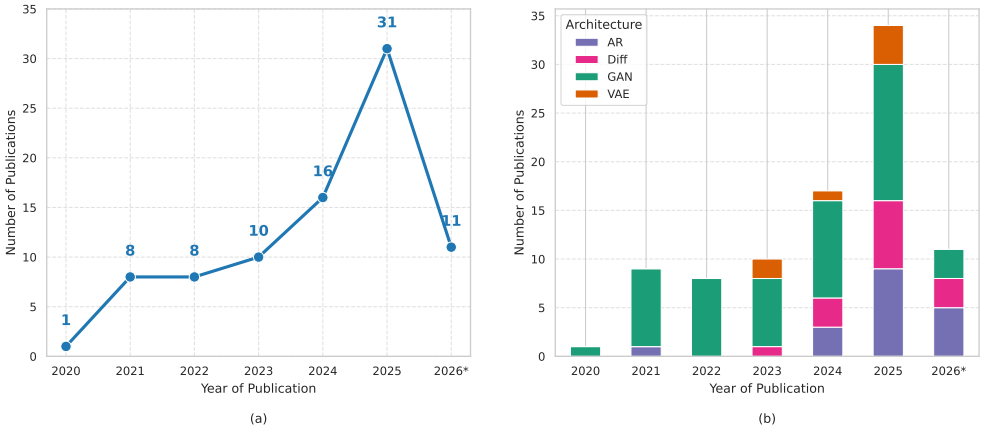


Fig. 3: Literature landscape and architectural evolution. (a) Overall publication trend per year; absolute publication counts are reported above each data point. (b) Yearly breakdown of the specific generative deep learning architectures adopted across the surveyed studies; note that individual articles may report multiple experiments using different architectures. The asterisk (*) on the year 2026 denotes that the data for the current year are partial and include publications collected only up to March.

To provide an overview of the growing interest and technical evolution in the field, we analyzed the annual distribution of the selected studies along with the specific generative architectures employed. As illustrated in Figure 3a, there is a clear upward trend in the number of articles published over time, peaking in 2025. Notably, with 12 publications recorded in just the first three months of 2026, the current year is on track to surpass the 2025 peak, reflecting the sustained momentum and expanding scope of research in this domain. In Figure 3b, the architectural breakdown reveals a noticeable paradigm shift: although GANs remain the most widely represented architecture across the surveyed literature, their relative dominance has gradually decreased. Conversely, architectures such as Diffusion models and AR frameworks, once representing a minor fraction, have gained significant traction, reflecting a broader diversification of generative techniques tailored to complex healthcare tasks.

The prevalence of GANs reflects a skewed distribution of modalities in FL applications, closely tied to the specific architectural requirements of generative models. As illustrated in Figure 4, image data are by far the most prevalent modality among the surveyed studies ($n = 58$), followed by multimodal ($n = 14$), sequential ($n = 11$), and tabular data ($n = 7$). This heavy bias aligns with the strong focus on medical imaging applications, which find particularly fertile ground in GANs due to their well-established maturity and key advantages in latent space manipulation and semantic editing [174]. Similarly, diffusion models display a strong specificity toward the image domain compared to other data modalities; notably, a growing body of recent literature is adopting diffusion-based architectures to tackle image-centric federated tasks.

In recent years, however, the body of work developing frameworks for tabular, sequential, and cross-modality data has increased, as illustrated in Figure 5, where non-imaging and multimodal modalities display a gradual upward trend over time. While partly due to the overall surge in research activity within the field, it also reflects a targeted effort to address specific clinical needs beyond medical imaging, facilitated

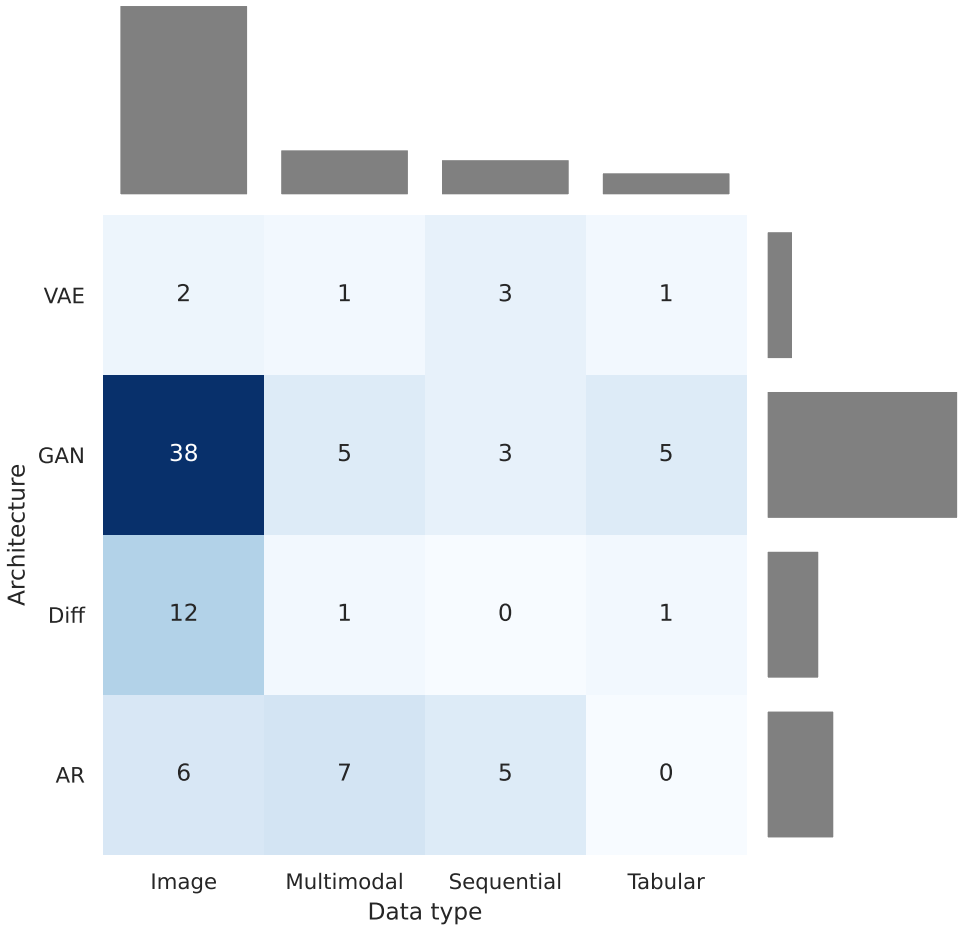


Fig. 4: The central heatmap displays the frequency of co-occurrence between different model architectures (e.g., GAN, VAE) and data types (e.g., Image, Tabular). Marginal histograms summarize the overall usage frequency of each architecture (top) and each data type (right). The color gradients in the central histograms reflect the relative intensity of usage. This visualization highlights common pairings and reveals preferences or gaps in current research practices.

by the inherent suitability of AR architectures for processing clinical sequences and reports, as well as their capacity to integrate heterogeneous data types for complex multimodal tasks. While AR accounted for less than 10% of surveyed studies before 2025, it has rapidly gained traction over the past two years, outpacing the adoption

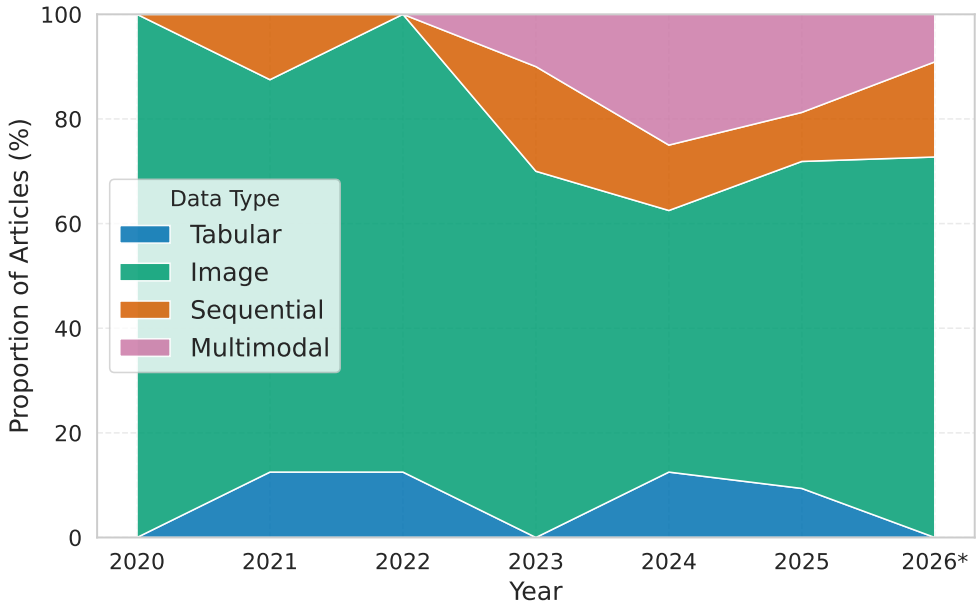


Fig. 5: Relative percentage of articles by data type over time. The visualization highlights the evolution of data type adoption in federated generative modeling for healthcare. The asterisk (*) on the year 2026 denotes that the data for the current year are partial and include publications collected only up to March.

of Diffusion models (see Figure 4). This shift reflects the broader trajectory of modern machine learning and healthcare applications, where transformers and foundation models have firmly established themselves as the prevailing paradigm [175, 176]. AR models are currently mainly used for text generation (report generation, interpretable diagnosis, clinical decision support) from multimodal sources, but AR transformers are also a natural fit for longitudinal EHRs, and only one paper currently explores this possibility [101].

Furthermore, a general analysis of the reviewed literature reveals an important observation: the consistent adoption of a centralized FL topology, where a central server coordinates the training process. Notably, in their work, Nguyen et al. [120] introduce the only example of decentralized design in the field of healthcare applications, with a blockchain-based peer-to-peer FL to enhance protection against attacks

on the central node. Although healthcare applications have largely relied on centralized FL topologies, recent research is beginning to shift the paradigm through tailored aggregation methods. Among the surveyed literature, some works are particularly notable for benchmarking alternative aggregation paradigms [88, 103, 117, 155, 168]. Moreover, beyond standard weight updates, innovative frameworks exchange latent representations or directly generated synthetic samples to synchronize knowledge across nodes [147, 148, 152, 153, 156].

6.1 Key Methodological and System Bottlenecks

While publication analysis underscores the main directions of research and trending questions, a critical synthesis of the underlying methodologies reveals several persistent architectural and technical challenges that hinder practical clinical deployment. Primary among these is the pervasive non-IID nature of health data across institutions, which severely destabilizes federated generative convergence. When local generative architectures fit site-specific biases, standard aggregation algorithms like FedAvg induce severe client drift and conflicting gradient updates, leading to degraded global representation quality or outright loss of divergence. This instability is particularly acute in adversarial setups such as FedGANs, where local discriminators overfit to institutional data boundaries, frequently triggering severe mode collapse [125, 146, 149]. To preserve global model stability across heterogeneous data silos, recent frameworks are increasingly moving toward personalized federated paradigms and adaptive, distance-weighted aggregation mechanisms.

Beyond data heterogeneity, the steep communication and computational demands of scaling high-capacity architectures, such as latent diffusion models and large AR transformers, impose severe constraints on bandwidth-limited healthcare networks. To ensure resource-constrained clinical nodes remain viable participants, current implementations progressively rely on parameter-efficient fine-tuning (PEFT) strategies

like LoRA, alongside gradient compression, latent space quantization, and federated knowledge distillation to exchange compact updates rather than full parameter matrices. Finally, despite clinical workflows being fundamentally multimodal, federated multimodal generative modeling remains disproportionately underrepresented in the reviewed literature. This systemic gap stems from substantial technical roadblocks, including the inability to maintain joint cross-modal alignment across decoupled client nodes, gradient dominance by high-dimensional spatial modalities during local training, and severe client drift induced by missing or incomplete modal streams across participating health systems [82, 84].

These findings provide a comprehensive overview of the current landscape of DGMs in healthcare FL, capturing a field that is currently transitioning toward greater architectural diversification. They highlight both the strengths, such as the widespread adoption of GANs and the focus on image data, and the gaps, including the significant hurdles in multimodal data integration and the lack of well-established and validated standards in FL. Addressing these gaps will be crucial for advancing the field and ensuring the practical applicability of these models in real-world healthcare settings.

7 Conclusion and future lines

This review systematically examined the integration of DGMs within FL frameworks in healthcare, selecting $n = 85$ articles published between 2020 and March 2026. The empirical analysis confirms an accelerating momentum in this domain, with publication counts rising from 1 in 2020 to 31 in 2025, and 12 recorded in the first quarter of 2026 alone. Methodologically, GANs maintain overall historical dominance with 51 appearances across surveyed tasks, comprising 66% of medical image synthesis studies, driven by their maturity in spatial feature modeling. Image data represent the vast majority of target modalities ($n = 58$), far outpacing multimodal ($n = 14$), sequential

($n = 11$), and tabular ($n = 7$) datasets. However, the field is undergoing a clear architectural shift toward multimodal diversification and foundation modeling. Diffusion models ($n = 14$) have rapidly emerged as a key approach for high-fidelity image generation, while AR models ($n = 18$) have expanded faster than any other family, rising from less than 10% of total studies prior to 2025 to become the leading choice for vision-language tasks and clinical report generation. Conversely, VAEs remain underrepresented, with only 7 studies. Despite multimodal frameworks emerging over the last three years ($n = 14$), joint cross-modal synthesis and longitudinal trajectory modeling remain largely under-explored. To transition these privacy-preserving distributed models into clinical practice, future efforts must establish standardized evaluation benchmarks across fidelity, utility, and explicit defenses against privacy leakage.

Looking forward, future research should concretely focus on establishing standardized evaluation benchmarks and comparative studies across multiple DGM architectures. Regulatory frameworks concerning data privacy, security, and compliance should be explicitly considered in future methodological developments, shaping how DGMs and FL systems can be safely and consistently deployed in clinical settings. Furthermore, researchers should actively explore novel architectures and hybrid combinations to establish robust cross-modality frameworks. These systems should be designed to integrate every available data modality to facilitate joint synthesis and causal exploration across biological systems by leveraging the synergy between both intra- and inter-modality dependencies, which is currently a notable gap in the literature. Equally important is the development of comprehensive metrics and evaluation protocols for data leakage. Future research must address not only the security of the training process against malicious updates but also the risk of inadvertent memorization of sensitive training data within the generative models themselves. Evaluation strategies must evolve beyond basic statistical similarity to include rigorous testing against leakage across the whole network. Ensuring that sensitive information does

not persist within the model parameters or the generated synthetic samples is essential for achieving the trustworthy and regulatory-compliant deployment required in real-world clinical environments.

Addressing these areas will not only enrich the landscape of DGMs in healthcare FL but also pave the way for compliant, secure, and clinically impactful generative solutions.

Acknowledgements. This work was supported by Synthetic Data Generation Framework for Integrated Validation of Use Cases and AI Healthcare Applications (SYNTHIA) project from Horizon Europe under Grant 101172872. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Declarations

- Funding: This work was supported by Synthetic Data Generation Framework for Integrated Validation of Use Cases and AI Healthcare Applications (SYNTHIA) project from Horizon Europe under Grant 101172872. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.
- Conflict of interest/Competing interests: The authors declare no competing interests.
- Ethics approval and consent to participate: This work is based solely on published literature and does not involve human participants or animals. As such, ethical approval and informed consent were not required.
- Consent for publication: Not applicable.

- Data availability: All data supporting the findings of this review are contained within the article and its tables. An interactive and extended version of the overview table is publicly available at https://gatv.gitlab.io/synthia/federated-gm-in-healthcare-review/.
- Materials availability: Not applicable.
- Code availability: The source code of the interactive table is available at https://gitlab.com/GATV/synthia/federated-gm-in-healthcare-review
- Author contribution: **Alessandro Ceresi:** Conceptualization, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review and editing, Supervision, Visualization. **Borja Arroyo Galende:** Conceptualization, Investigation, Data curation, Writing – original draft, Writing – review and editing. **Javier Guinea-Pérez:** Conceptualization, Data Curation, Investigation, Writing - original draft, Writing - review and editing, Visualization. **Patricia A. Apellániz:** Conceptualization, Writing - Original Draft, Writing - Review & Editing, Supervision. **Gustavo Hernández-Peñaloza:** Conceptualization, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. **Federico Álvarez:** Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

References

[1] Bekbolatova M, Mayer J, Ong CW, et al (2024) Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. Healthcare 12(2):125. https://doi.org/10.3390/healthcare12020125, number: 2

[2] Al Kuwaiti A, Nazer K, Al-Reedy A, et al (2023) A Review of the Role of Artificial Intelligence in Healthcare. Journal of Personalized Medicine 13(6):951. https://doi.org/10.3390/jpm13060951, number: 6

[3] Mullankandy S, Mukherjee S, Ingole BS (2024) Applications of AI in Electronic Health Records, Challenges, and Mitigation Strategies. In: 2024 International Conference on Computer and Applications (ICCA), pp 1–7, https://doi.org/10.1109/ICCA62237.2024.10927863

[4] Castiglioni I, Rundo L, Codari M, et al (2021) AI applications to medical images: From machine learning to deep learning. Physica Medica 83:9–24. https://doi.org/10.1016/j.ejmp.2021.02.006

[5] Subhan F, Mirza A, Su'ud MBM, et al (2023) AI-Enabled Wearable Medical Internet of Things in Healthcare System: A Survey. Applied Sciences 13(3):1394. https://doi.org/10.3390/app13031394, number: 3

[6] Williamson SM, Prybutok V (2024) Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. Applied Sciences 14(2):675. https://doi.org/10.3390/app14020675, number: 2

[7] Wojtara M, Rana E, Rahman T, et al (2023) Artificial intelligence in rare disease diagnosis and treatment. Clinical and Translational Science 16(11):2106–2111. https://doi.org/10.1111/cts.13619

[8] Zhang P, Kamel Boulos MN (2023) Generative AI in Medicine and Healthcare: Promises, Opportunities and Challenges. Future Internet 15(9):286. https://doi.org/10.3390/fi15090286, number: 9

[9] Frid-Adar M, Klang E, Amitai M, et al (2018) Synthetic data augmentation using GAN for improved liver lesion classification. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp 289–293, https://doi.org/10.1109/ISBI.2018.8363576

[10] Yoon J, Jordon J, Schaar M (2018) GAIN: Missing Data Imputation using Generative Adversarial Nets. In: Proceedings of the 35th International Conference on Machine Learning. PMLR, pp 5689–5698, URL https://proceedings.mlr.press/v80/yoon18a.html

[11] Baowaly MK, Lin CC, Liu CL, et al (2019) Synthesizing electronic health records using improved generative adversarial networks. Journal of the American Medical Informatics Association: JAMIA 26(3):228–241. https://doi.org/10.1093/jamia/ocy142

[12] Li J, Cairns BJ, Li J, et al (2023) Generating synthetic mixed-type longitudinal electronic health records for artificial intelligent applications. npj Digital Medicine 6(1):1–18. https://doi.org/10.1038/s41746-023-00834-7

[13] European Union (2016) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). URL https://eur-lex.europa.eu/eli/reg/2016/679/oj

[14] Directorate-General for Health and Food Safety (DG SANTE), European Commission (2025) Regulation (EU) 2025/327 of the European Parliament and of the Council on the European Health Data Space. URL https://eur-lex.europa.eu/eli/reg/2025/327/oj

[15] U.S. Department of Health and Human Services, Office for Civil Rights (2025) Notice of Proposed Rulemaking to Update the HIPAA Security Rule. URL https://www.hhs.gov/hipaa/for-professionals/security/hipaa-security-rule-nprm/index.html

[16] McMahan B, Moore E, Ramage D, et al (2017) Communication-Efficient Learning of Deep Networks from Decentralized Data. In: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. PMLR, pp 1273–1282, URL https://proceedings.mlr.press/v54/mcmahan17a.html

[17] Bonawitz K, Ivanov V, Kreuter B, et al (2017) Practical Secure Aggregation for Privacy-Preserving Machine Learning. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. Association for Computing Machinery, New York, NY, USA, CCS '17, pp 1175–1191, https://doi.org/10.1145/3133956.3133982

[18] Ogburn M, Turner C, Dahal P (2013) Homomorphic Encryption. Procedia Computer Science 20:502–509. https://doi.org/10.1016/j.procs.2013.09.310

[19] Goodfellow IJ, Pouget-Abadie J, Mirza M, et al (2014) Generative Adversarial Nets. In: Advances in Neural Information Processing Systems, vol 27. Curran Associates, Inc., URL https://proceedings.neurips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html

[20] Kingma DP, Welling M (2022) Auto-Encoding Variational Bayes. https://doi.org/10.48550/arXiv.1312.6114, arXiv:1312.6114 [stat.ML]

[21] Choi E, Biswal S, Malin B, et al (2017) Generating Multi-label Discrete Patient Records using Generative Adversarial Networks. In: Proceedings of the 2nd Machine Learning for Healthcare Conference. PMLR, pp 286–305, URL https://proceedings.mlr.press/v68/choi17a.html

[22] Sai S, Gaur A, Sai R, et al (2024) Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies, and Limitations. IEEE Access 12:31078–31106. https://doi.org/10.1109/ACCESS.2024.3367715

[23] Gm H, Gourisaria MK, Pandey M, et al (2020) A comprehensive survey and analysis of generative models in machine learning. Computer Science Review 38:100285. https://doi.org/10.1016/j.cosrev.2020.100285

[24] Lamb A (2021) A Brief Introduction to Generative Models. URL https://arxiv.org/abs/2103.00265v1

[25] Ou Z (2018) A Review of Learning with Deep Generative Models from Perspective of Graphical Modeling. https://doi.org/10.48550/ARXIV.1808.01630, version Number: 4

[26] White T (2016) Sampling Generative Networks. https://doi.org/10.48550/arXiv.1609.04468, arXiv:1609.04468

[27] Ho J, Jain A, Abbeel P (2020) Denoising Diffusion Probabilistic Models. https://doi.org/10.48550/arXiv.2006.11239, arXiv:2006.11239

[28] Rezende DJ, Mohamed S (2016) Variational Inference with Normalizing Flows. https://doi.org/10.48550/arXiv.1505.05770, arXiv:1505.05770

[29] Ackley DH, Hinton GE, Sejnowski TJ (1985) A learning algorithm for boltzmann machines. Cognitive Science 9(1):147–169. https://doi.org/10.1016/S0364-0213(85)80012-4

[30] Challen R, Denny J, Pitt M, et al (2019) Artificial intelligence, bias and clinical safety. BMJ Quality & Safety 28(3):231–237. https://doi.org/10.1136/bmjqs-2018-008370

[31] Frid-Adar M, Diamant I, Klang E, et al (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 321:321–331. https://doi.org/10.1016/j.neucom.2018.09.013

[32] Chen RJ, Lu MY, Chen TY, et al (2021) Synthetic data in machine learning for medicine and healthcare. Nature Biomedical Engineering 5(6):493–497. https://doi.org/10.1038/s41551-021-00751-8

[33] Yoon J, Jarrett D, van der Schaar M (2019) Time-series Generative Adversarial Networks. In: Advances in Neural Information Processing Systems, vol 32. Curran Associates, Inc., URL https://papers.nips.cc/paper_files/paper/2019/hash/c9efe5f26cd17ba6216bbe2a7d26d490-Abstract.html

[34] Beaulieu-Jones B, Wu Z, Williams C, et al (2019) Privacy-Preserving Generative Deep Neural Networks Support Clinical Data Sharing. Circulation: Cardiovascular Quality and Outcomes 12. https://doi.org/10.1161/CIRCOUTCOMES.118.005122

[35] Razavian N, Marcus J, Sontag D (2016) Multi-task Prediction of Disease Onsets from Longitudinal Laboratory Tests. In: Proceedings of the 1st Machine Learning for Healthcare Conference. PMLR, pp 73–100, URL https://proceedings.mlr.press/v56/Razavian16.html

[36] Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine 25(1):44–56. https://doi.org/10.1038/s41591-018-0300-7

[37] (2024) Introducing ChatGPT. URL https://openai.com/index/chatgpt/

[38] Dalal M, Li AC, Taori R (2019) Autoregressive Models: What Are They Good For? URL https://arxiv.org/abs/1910.07737v1

[39] Müller-Franzes G, Niehues JM, Khader F, et al (2023) A multimodal comparison of latent denoising diffusion probabilistic models and generative adversarial networks for medical image synthesis. Scientific Reports 13(1):12098. https:

//doi.org/10.1038/s41598-023-39278-0

[40] Zhong Y, Wang X, Wang J, et al (2024) Synthesizing Multimodal Electronic Health Records via Predictive Diffusion Models. https://doi.org/10.48550/arXiv.2406.13942, arXiv:2406.13942 [cs]

[41] Majeed IA, Kaushik S, Bardhan A, et al (2022) Comparative assessment of federated and centralized machine learning. https://doi.org/10.48550/arXiv.2202.01529, arXiv:2202.01529

[42] Yang Q, Liu Y, Chen T, et al (2019) Federated Machine Learning: Concept and Applications. ACM Trans Intell Syst Technol 10(2):12:1–12:19. https://doi.org/10.1145/3298981

[43] Ahn S, Kim S, Kwon Y, et al (2023) Communication-Efficient Diffusion Strategy for Performance Improvement of Federated Learning with Non-IID Data. https://doi.org/10.48550/arXiv.2207.07493, arXiv:2207.07493 [cs]

[44] Mammen PM (2021) Federated Learning: Opportunities and Challenges. https://doi.org/10.48550/arXiv.2101.05428, arXiv:2101.05428 [cs]

[45] Truong N, Sun K, Wang S, et al (2021) Privacy preservation in federated learning: An insightful survey from the GDPR perspective. Computers & Security 110:102402. https://doi.org/10.1016/j.cose.2021.102402

[46] Gostin LO (2001) National health information privacy: regulations under the Health Insurance Portability and Accountability Act. JAMA 285(23):3015–3021. https://doi.org/10.1001/jama.285.23.3015

[47] Ji Z, Lipton ZC, Elkan C (2014) Differential Privacy and Machine Learning: a Survey and Review. https://doi.org/10.48550/arXiv.1412.7584, arXiv:1412.7584

[48] Lindell Y (2021) Secure multiparty computation. Communications of the ACM 64(1):86–96. https://doi.org/10.1145/3387108

[49] Li T, Sahu AK, Zaheer M, et al (2018) Federated Optimization in Heterogeneous Networks. URL https://arxiv.org/abs/1812.06127v5

[50] Zhang C, Xie Y, Bai H, et al (2021) A survey on federated learning. Knowledge-Based Systems 216:106775. https://doi.org/10.1016/j.knosys.2021.106775

[51] Rujas M, Martín Gómez del Moral Herranz R, Fico G, et al (2025) Synthetic data generation in healthcare: A scoping review of reviews on domains, motivations, and future applications. International Journal of Medical Informatics 195:105763. https://doi.org/10.1016/j.ijmedinf.2024.105763

[52] Thanh-Tung H, Tran T (2021) Toward a Generalization Metric for Deep Generative Models. https://doi.org/10.48550/arXiv.2011.00754, arXiv:2011.00754 [cs]

[53] Regenwetter L, Srivastava A, Gutfreund D, et al (2023) Beyond Statistical Similarity: Rethinking Metrics for Deep Generative Models in Engineering Design. Computer-Aided Design 165:103609. https://doi.org/10.1016/j.cad.2023.103609

[54] Chen N, Klushyn A, Kurle R, et al (2018) Metrics for Deep Generative Models. In: Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics. PMLR, pp 1540–1550, URL https://proceedings.mlr.press/v84/chen18e.html

[55] Palacio-Niño JO, Berzal F (2019) Evaluation Metrics for Unsupervised Learning Algorithms. URL https://arxiv.org/abs/1905.05667v2

[56] Boursalie O, Samavi R, Doyle TE (2022) Evaluation Metrics for Deep Learning Imputation Models. In: Shaban-Nejad A, Michalowski M, Bianco S (eds) AI for Disease Surveillance and Pandemic Intelligence: Intelligent Disease Detection in Action. Springer International Publishing, Cham, p 309–322, https://doi.org/10.1007/978-3-030-93080-6_22

[57] Hernandez M, Epelde G, Alberdi A, et al (2022) Synthetic data generation for tabular health records: A systematic review. Neurocomputing 493:28–45. https://doi.org/10.1016/j.neucom.2022.04.053

[58] Figueira A, Vaz B (2022) Survey on Synthetic Data Generation, Evaluation Methods and GANs. Mathematics 10(15):2733. https://doi.org/10.3390/math10152733, number: 15

[59] Mendelevitch O, Lesh MD (2021) Fidelity and Privacy of Synthetic Medical Data. https://doi.org/10.48550/arXiv.2101.08658, arXiv:2101.08658

[60] El Emam K (2020) Seven Ways to Evaluate the Utility of Synthetic Data. IEEE Security & Privacy 18(4):56–59. https://doi.org/10.1109/MSEC.2020.2992821, conference Name: IEEE Security & Privacy

[61] Galende BA, Apellániz PA, Parras J, et al (2025) Membership Inference Attacks and Differential Privacy: A Study Within the Context of Generative Models. IEEE Open Journal of the Computer Society 6:801–811. https://doi.org/10.1109/OJCS.2025.3572244

[62] (2024) Executive Summary for the Digital Health Advisory Committee Meeting: Total Product Lifecycle Considerations for Generative AI-Enabled Devices. Executive Summary, U.S. Food & Drug Administration, Silver Spring, MD, URL https://www.fda.gov/media/174897/download

[63] European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). URL https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[64] Hayes J, Melis L, Danezis G, et al (2019) LOGAN: Membership Inference Attacks Against Generative Models. Proceedings on Privacy Enhancing Technologies 2019(1):133–152. https://doi.org/10.2478/popets-2019-0008

[65] Abadi M, Chu A, Goodfellow I, et al (2016) Deep Learning with Differential Privacy. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. ACM, Vienna Austria, pp 308–318, https://doi.org/10.1145/2976749.2978318

[66] Bagdasaryan E, Veit A, Hua Y, et al (2020) How To Backdoor Federated Learning. In: Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics. PMLR, pp 2938–2948, URL https://proceedings.mlr.press/v108/bagdasaryan20a.html

[67] Hitaj B, Ateniese G, Perez-Cruz F (2017) Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, Dallas Texas USA, pp 603–618, https://doi.org/10.1145/3133956.3134012

[68] Tramèr F, Zhang F, Juels A, et al (2016) Stealing Machine Learning Models via Prediction {APIs}. pp 601–618, URL https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/tramer

[69] Berryman JG (1985) Measurement of spatial correlation functions using image processing techniques. Journal of Applied Physics 57(7):2374–2384. https://doi.

org/10.1063/1.334346

[70] Alvarez-Lacalle E, Dorow B, Eckmann JP, et al (2006) Hierarchical structures induce long-range dynamical correlations in written texts. Proceedings of the National Academy of Sciences 103(21):7956–7961. https://doi.org/10.1073/pnas.0510673103

[71] Patel R, Verma AK, Sahu NK, et al (2025) Generative AI Techniques in Medical Imaging Analysis: A Systematic Review. In: Real-Time Data Decisions With AI and ChatGPT Techniques. IGI Global, p 91–126, https://doi.org/10.4018/979-8-3693-2284-0.ch005

[72] Kaissis GA, Makowski MR, Rückert D, et al (2020) Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence 2(6):305–311. https://doi.org/10.1038/s42256-020-0186-1

[73] Bartell JA, Valentin SB, Krogh A, et al (2024) A primer on synthetic health data. URL http://arxiv.org/abs/2401.17653, arXiv:2401.17653

[74] Kang HYJ, Ko M, Ryu KS (2024) Tabular Transformer Generative Adversarial Network for Heterogeneous distribution in healthcare. https://doi.org/10.21203/rs.3.rs-4134206/v1, iSSN: 2693-5015

[75] Ghosheh GO, Thwaites CL, Zhu T (2023) Synthesizing Electronic Health Records for Predictive Models in Low-Middle-Income Countries (LMICs). Biomedicines 11(6):1749. https://doi.org/10.3390/biomedicines11061749, number: 6

[76] Ma C, Tschiatschek S, Palla K, et al (2019) EDDI: Efficient Dynamic Discovery of High-Value Information with Partial VAE. https://doi.org/10.48550/arXiv.1809.11142, arXiv:1809.11142 [cs]

[77] Yang H, Ge M, Xiang K, et al (2023) FedVAE: Communication-Efficient Federated Learning With Non-IID Private Data. IEEE Systems Journal 17(3):4798–4808. https://doi.org/10.1109/JSYST.2023.3274197, conference Name: IEEE Systems Journal

[78] Häyrinen K, Saranto K, Nykänen P (2008) Definition, structure, content, use and impacts of electronic health records: A review of the research literature. International Journal of Medical Informatics 77(5):291–304. https://doi.org/10.1016/j.ijmedinf.2007.09.001

[79] Evans RS (2018) Electronic Health Records: Then, Now, and in the Future. Yearbook of Medical Informatics 25:S48–S61. https://doi.org/10.15265/IYS-2016-s006

[80] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539

[81] Zhang Y, Chen M, Shen J, et al (2022) Tailor Versatile Multi-Modal Learning for Multi-Label Emotion Recognition. Proceedings of the AAAI Conference on Artificial Intelligence 36(8):9100–9108. https://doi.org/10.1609/aaai.v36i8.20895, number: 8

[82] Dimitri GM (2022) A Short Survey on Deep Learning for Multimodal Integration: Applications, Future Perspectives and Challenges. Computers 11(11):163. https://doi.org/10.3390/computers11110163, number: 11

[83] Baltrušaitis T, Ahuja C, Morency LP (2018) Challenges and applications in multimodal machine learning. In: The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition - Volume 2, vol 21. Association for Computing Machinery and Morgan &

Claypool, p 17–48, URL https://doi.org/10.1145/3107990.3107993

[84] Lipkova J, Chen RJ, Chen B, et al (2022) Artificial intelligence for multimodal data integration in oncology. Cancer Cell 40(10):1095–1110. https://doi.org/10.1016/j.ccell.2022.09.012

[85] Cui C, Yang H, Wang Y, et al (2023) Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review. Progress in Biomedical Engineering 5(2):022001. https://doi.org/10.1088/2516-1091/acc2fe

[86] Lorenzi M, Deprez M, Balelli I, et al (2023) Integration of Multimodal Data. In: Colliot O (ed) Machine Learning for Brain Disorders. Springer US, New York, NY, p 573–597, https://doi.org/10.1007/978-1-0716-3195-9_19

[87] Serrano ALM, Rodrigues GAP, Bispo GD, et al (2026) FedIHRAS: A Privacy-Preserving Federated Learning Framework for Multi-Institutional Collaborative Radiological Analysis with Integrated Explainability and Automated Clinical Reporting. BIOMEDICINES 14(3). https://doi.org/10.3390/biomedicines14030713

[88] Metmer H, Yang X (2025) FedMRG: federated medical report generation via text-aware learning rate adjustment and multi-level prototype collaboration. MULTIMEDIA SYSTEMS 31(2). https://doi.org/10.1007/s00530-025-01725-5

[89] Che H, Jin H, Guo Z, et al (2026) LLM-Driven Medical Report Generation via Communication-Efficient Heterogeneous Federated Learning. IEEE TRANSACTIONS ON MEDICAL IMAGING 45(1):28–39. https://doi.org/10.1109/TMI.2025.3591185

[90] Zhu H, Togo R, Ogawa T, et al (2026) Personalized federated learning for medical vision-language models via efficient fine-tuning and uncertainty-aware

disentanglement. Journal of biomedical informatics 178:105014–105014. https://doi.org/10.1016/j.jbi.2026.105014

[91] Hossain M, Ahmed M, Samu M, et al (2025) Privacy-Preserving Chest X-Ray Report Generation via Multimodal Federated Learning with ViT and GPT2. Biomedical Materials and Devices https://doi.org/10.1007/s44174-025-00538-4

[92] Ghanta S, Thiriveedhi A, Boyapati P, et al (2025) Federated Transfer Learning for Chest X-ray Classification: An Explainable and Generative AI Framework with Reliability Assessment. SN Computer Science 6(7). https://doi.org/10.1007/s42979-025-04315-4

[93] Yan B, Chen Q, Chen Y, et al (2024) Buffalo: Biomedical Vision-Language Understanding with Cross-Modal Prototype and Federated Foundation Model Collaboration. In: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. Association for Computing Machinery, New York, NY, USA, CIKM '24, pp 2775–2785, https://doi.org/10.1145/3627673.3679627

[94] Kumar R, Rao Y, Kumar J, et al (2025) MID-LLM: Enhancing Medical Image Diagnostics With LLMs in a Blockchain AI Framework. IEEE INTERNET OF THINGS JOURNAL 12(18):37160–37174. https://doi.org/10.1109/JIOT.2025.3583408

[95] Chen J, Pan R (2024) Medical report generation based on multimodal federated learning. Computerized Medical Imaging and Graphics 113:102342. https://doi.org/10.1016/j.compmedimag.2024.102342

[96] Mukhopadhyay S, Haider N, Chakraborty C, et al (2026) GeoFed-Cervix: A Differential Geometry-Guided Federated and Explainable AI Framework for Early Cervical Cancer Detection on Consumer Devices. IEEE TRANSACTIONS ON

CONSUMER ELECTRONICS 72(1):2360–2367. https://doi.org/10.1109/TCE.2026.3651019

[97] Karamanlioglu A, Demirel B, Tural O, et al (2025) Privacy-Preserving Clinical Decision Support for Emergency Triage Using LLMs: System Architecture and Real-World Evaluation. APPLIED SCIENCES-BASEL 15(15). https://doi.org/10.3390/app15158412

[98] Ben Othman S, Ali O (2025) Quantum-entangled neuro-symbolic swarm federation for privacy-preserving IoMT-driven multimodal healthcare. SCIENTIFIC REPORTS 15(1). https://doi.org/10.1038/s41598-025-31820-6

[99] Hashjin NM, Amiri MH, Najafabadi MK (2026) DermaGPT a federated multimodal framework with a meta learned trust function for interpretable dermatology diagnostics. SCIENTIFIC REPORTS 16(1). https://doi.org/10.1038/s41598-026-38715-0

[100] Puppala S, Hossain I, Alam MJ, et al (2024) SCAN: A HealthCare Personalized ChatBot with Federated Learning Based GPT. pp 1945–1951, https://doi.org/10.1109/COMPSAC61105.2024.00309

[101] Kim S, Lee KJ, Kim T, et al (2025) Balancing Privacy and Accuracy in Healthcare AI: Federated Learning with AutoML for Blood Pressure Prediction. APPLIED SCIENCES-BASEL 15(19). https://doi.org/10.3390/app151910624

[102] Li Z, Dib O, Li S (2026) Data-centric federated learning for neuro-oncology: Addressing heterogeneity via privacy-preserving generative augmentation. ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE 173. https://doi.org/10.1016/j.engappai.2026.114351

[103] Zhou Z, Luo G, Chen M, et al (2025) Federated Learning for Medical Image Classification: A Comprehensive Benchmark. IEEE journal of biomedical and health informatics https://doi.org/10.1109/JBHI.2025.3631706

[104] Niu K, Cai J, Feng X (2026) SAFedHDM: Semi-asynchronous federated learning with highlight diffusion model for medical image segmentation. INFORMATION FUSION 126. https://doi.org/10.1016/j.inffus.2025.103615

[105] K. Huangsuwan, T. Liu, S. See, et al (2025) FedDrip: Federated Learning With Diffusion-Generated Synthetic Image. IEEE Access 13:10111–10125. https://doi.org/10.1109/ACCESS.2025.3525806

[106] H. Sun, W. Zhang, R. Zhang, et al (2025) Federated Continual Learning Based on Weakly Supervised Diffusion Models for Disease Diagnosis. IEEE Internet of Things Journal 12(22):45958–45971. https://doi.org/10.1109/JIOT.2025.3535628

[107] Zhang H, Chen M, Liu Y, et al (2025) Non-IID Medical Image Segmentation Based on Cascaded Diffusion Model for Diverse Multi-Center Scenarios. IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS 29(7):5042–5055. https://doi.org/10.1109/JBHI.2025.3549029

[108] Yoon T, Hwang M, Yang E (2025) VQ-FedDiff: Federated Learning Algorithm of Diffusion Models With Client-Specific Vector-Quantized Conditioning. IEEE Transactions on Pattern Analysis and Machine Intelligence 47(12):11863–11873. https://doi.org/10.1109/TPAMI.2025.3602282

[109] Song X, Shu K, Yang P, et al (2026) Developing a knowledge-guided federated graph attention learning network with a diffusion module to diagnose Alzheimer's disease. MEDICAL IMAGE ANALYSIS 107. https://doi.org/10.

1016/j.media.2025.103794

[110] Tun YL, Thwal CM, Yoon JS, et al (2023) Federated Learning with Diffusion Models for Privacy-Sensitive Vision Tasks. In: 2023 International Conference on Advanced Technologies for Communications (ATC), pp 305–310, https://doi.org/10.1109/ATC58710.2023.10318858

[111] Li Z, Kainz B (2024) Image Distillation for Safe Data Sharing in Histopathology. In: Linguraru MG, Dou Q, Feragen A, et al (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. Springer Nature Switzerland, Cham, pp 459–469, https://doi.org/10.1007/978-3-031-72117-5_43

[112] Wu L, Liu Z, Pu B, et al (2025) DGGI: Deep Generative Gradient Inversion with diffusion model. Information Fusion 113. https://doi.org/10.1016/j.inffus.2024.102620

[113] Sattarov T, Schreyer M, Borth D (2024) FedTabDiff: Federated Learning of Diffusion Probabilistic Models for Synthetic Mixed-Type Tabular Data Generation. https://doi.org/10.48550/arXiv.2401.06263, arXiv:2401.06263 [cs]

[114] Hou H, Zhang J, Zhang M, et al (2025) BiFDR: Brain-Inspired Federated Diffusion Transformer with Reinforcement for privacy-preserving molecular generation. JOURNAL OF BIOMEDICAL INFORMATICS 170. https://doi.org/10.1016/j.jbi.2025.104910

[115] Pang M, Xu H, Huang Z, et al (2026) A Tri-Factor Adaptive Federated Learning Framework for Parkinson's Disease Diagnosis via Multi-Source Facial Expression Analysis. IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS 30(3):2500–2511. https://doi.org/10.1109/JBHI.2025.3605732

[116] Cai X, Lan Y, Zhang Z, et al (2023) A Many-Objective Optimization Based Federal Deep Generation Model for Enhancing Data Processing Capability in IoT. IEEE Transactions on Industrial Informatics 19(1):561–569. https://doi.org/10.1109/TII.2021.3093715

[117] Li Q, Jiang Y, Zhong R, et al (2025) FedG3FA: Three-Stage GAN-Aided Target Feature Alignment for Secure Data Sharing in Federated Learning System. IEEE Transactions on Information Forensics and Security 20:9806–9817. https://doi.org/10.1109/TIFS.2025.3609664

[118] Li S, Hu L, Sun C, et al (2025) Federated edge learning for medical image augmentation. Applied Intelligence 55(1). https://doi.org/10.1007/s10489-024-06046-0

[119] Han T, Nebelung S, Haarburger C, et al (2020) Breaking medical data sharing boundaries by using synthesized radiographs. Science Advances 6(49):eabb7973. https://doi.org/10.1126/sciadv.abb7973

[120] Nguyen DC, Ding M, Pathirana PN, et al (2022) Federated Learning for COVID-19 Detection With Generative Adversarial Networks in Edge Cloud Computing. IEEE Internet of Things Journal 9(12):10257–10271. https://doi.org/10.1109/JIOT.2021.3120998

[121] Liu W, Zheng Y, Xiang Z, et al (2025) An efficient federated learning method based on enhanced classification-GAN for medical image classification. Multimedia Systems 31(1). https://doi.org/10.1007/s00530-024-01564-w

[122] Biswas S, Mostafiz R, Uddin MS, et al (2025) FLPneXAINet: Federated deep learning and explainable AI for improved pneumonia prediction utilizing GAN-augmented chest X-ray data. PLOS ONE 20(7). https://doi.org/10.1371/

journal.pone.0324957

[123] Mahir S, Tashrif M, Shathe A, et al (2026) PriFL-XAI: Hybrid privacy-preserving federated learning models for monkeypox detection through GAN augmentation and explainable AI. Biomedical Signal Processing and Control 112. https://doi.org/10.1016/j.bspc.2025.108426

[124] Lan Y, Xie L, Cai X, et al (2022) A many-objective evolutionary algorithm based on integrated strategy for skin cancer detection. KSII Transactions on Internet and Information Systems (TIIS) 16(1):80–96. https://doi.org/10.3837/tiis.2022.01.005

[125] Jin R, Li X (2023) Backdoor attack and defense in federated generative adversarial network-based medical image synthesis. Medical Image Analysis 90:102965. https://doi.org/10.1016/j.media.2023.102965

[126] Alalwan N, Alwadain A, Alzahrani AI, et al (2024) Advancements in brain tumor identification: Integrating synthetic GANs with federated-CNNs in medical imaging analysis. Alexandria Engineering Journal 105:105–119. https://doi.org/10.1016/j.aej.2024.06.080

[127] Peng L, Wang N, Dvornek N, et al (2023) FedNI: Federated Graph Learning With Network Inpainting for Population-Based Disease Prediction. IEEE Transactions on Medical Imaging 42(7):2032–2043. https://doi.org/10.1109/TMI.2022.3188728

[128] Yang J, Gu Y, Gao S, et al (2025) Fed-GFM-DG: A privacy-preserving framework for fundus image segmentation via generative-based feature generalization and mask-guided aggregation. Biomedical Signal Processing and Control 106. https://doi.org/10.1016/j.bspc.2025.107725

[129] Peketi D, Chalavadi V, Mohan CK, et al (2023) FLWGAN: Federated Learning with Wasserstein Generative Adversarial Network for Brain Tumor Segmentation. In: 2023 International Joint Conference on Neural Networks (IJCNN), pp 1–8, https://doi.org/10.1109/IJCNN54540.2023.10191202

[130] Nimeshika GN, D S (2024) Enhancing Alzheimer's disease classification through split federated learning and GANs for imbalanced datasets. PeerJ Computer Science 10:e2459. https://doi.org/10.7717/peerj-cs.2459

[131] Zhang L, Shen B, Barnawi A, et al (2021) FedDPGAN: Federated Differentially Private Generative Adversarial Networks Framework for the Detection of COVID-19 Pneumonia. Information Systems Frontiers: A Journal of Research and Innovation 23(6):1403–1415. https://doi.org/10.1007/s10796-021-10144-6

[132] Lee EH, Kelly B, Altinmakas E, et al (2023) AI Models Close to your Chest: Robust Federated Learning Strategies for Multi-site CT. https://doi.org/10.48550/arXiv.2303.13567, arXiv:2303.13567 [cs]

[133] Abdel-Basset M, Hawash H, Abouhawwash M (2022) Collaborative Screening of COVID-19-like Disease from Multi-Institutional Radiographs: A Federated Learning Approach. Mathematics 10(24):4766. https://doi.org/10.3390/math10244766, number: 24

[134] Ahmed S, Feng J, Ferzund J, et al (2025) FAME: A Federated Adversarial Learning Framework for Privacy-Preserving MRI Reconstruction. Applied Magnetic Resonance 56(6):681–722. https://doi.org/10.1007/s00723-025-01749-0

[135] Kundu D, Rahman MM, Rahman A, et al (2024) Federated Deep Learning for Monkeypox Disease Detection on GAN-Augmented Dataset. IEEE Access 12:32819–32829. https://doi.org/10.1109/ACCESS.2024.3370838

[136] Izhar A, Anwar SM (2025) FL-Swarm MRCM: A Novel Federated Learning Framework for Cross-Site Medical Image Reconstruction. BIG DATA AND COGNITIVE COMPUTING 9(11). https://doi.org/10.3390/bdcc9110295

[137] Shen Y, Sowmya A, Luo Y, et al (2023) A Federated Learning System for Histopathology Image Analysis With an Orchestral Stain-Normalization GAN. IEEE Transactions on Medical Imaging 42(7):1969–1981. https://doi.org/10.1109/TMI.2022.3221724, conference Name: IEEE Transactions on Medical Imaging

[138] Murmu A, Kumar P, Namasudra S, et al (2025) GAN-Based Federated Adversarial Learning for Enhancing Security Towards Consumer Digital Ecosystems. IEEE Transactions on Consumer Electronics 71(1):1102–1114. https://doi.org/10.1109/TCE.2024.3522018

[139] Murmu A, Kumar P, Rao Moparthi N, et al (2024) Reliable Federated Learning With GAN Model for Robust and Resilient Future Healthcare System. IEEE Transactions on Network and Service Management 21(5):5335–5346. https://doi.org/10.1109/TNSM.2024.3422376

[140] Mazher M, Razzak I, Qayyum A, et al (2024) Self-supervised spatial–temporal transformer fusion based federated framework for 4D cardiovascular image segmentation. Information Fusion 106:102256. https://doi.org/10.1016/j.inffus.2024.102256

[141] Georgiadis A, Babbar V, Silavong F, et al (2022) ST-FL: style transfer preprocessing in federated learning for COVID-19 segmentation. In: Medical Imaging 2022: Imaging Informatics for Healthcare, Research, and Applications, vol 12037. SPIE, pp 13–27, https://doi.org/10.1117/12.2611096

[142] Yan Z, Wicaksana J, Wang Z, et al (2021) Variation-Aware Federated Learning With Multi-Source Decentralized Medical Image Data. IEEE Journal of Biomedical and Health Informatics 25(7):2615–2628. https://doi.org/10.1109/JBHI.2020.3040015

[143] Elmas G, Dar SUH, Korkmaz Y, et al (2023) Federated Learning of Generative Image Priors for MRI Reconstruction. IEEE Transactions on Medical Imaging 42(7):1996–2009. https://doi.org/10.1109/TMI.2022.3220757, conference Name: IEEE Transactions on Medical Imaging

[144] Wang J, Xie G, Huang Y, et al (2023) FedMed-GAN: Federated domain translation on unsupervised cross-modality brain image synthesis. Neurocomputing 546:126282. https://doi.org/10.1016/j.neucom.2023.126282

[145] Li Y, Tan Q, Shin BS (2025) CryptoGAN: Privacy-Preserving Federated Generative Adversarial Networks With Homomorphic Encryption in Healthcare Systems. IEEE Transactions on Computational Social Systems 12(6):5330–5341. https://doi.org/10.1109/TCSS.2025.3570990

[146] Kamran H, Hussain S, Latif S, et al (2025) FedGAN: Federated diabetic retinopathy image generation. PLOS ONE 20(7). https://doi.org/10.1371/journal.pone.0326579

[147] Pennisi M, Proietto Salanitri F, Palazzo S, et al (2022) GAN Latent Space Manipulation and Aggregation for Federated Learning in Medical Imaging. In: Albarqouni S, Bakas S, Bano S, et al (eds) Distributed, Collaborative, and Federated Learning, and Affordable AI and Healthcare for Resource Diverse Global Health. Springer Nature Switzerland, Cham, pp 68–78, https://doi.org/10.1007/978-3-031-18523-6_7

[148] Rajotte JF, Mukherjee S, Robinson C, et al (2021) Reducing bias and increasing utility by federated generative modeling of medical images using a centralized adversary. In: Proceedings of the Conference on Information Technology for Social Good. Association for Computing Machinery, New York, NY, USA, GoodIT '21, pp 79–84, https://doi.org/10.1145/3462203.3475875

[149] Jin R, Li X (2022) Backdoor Attack is a Devil in Federated GAN-Based Medical Image Synthesis. In: Zhao C, Svoboda D, Wolterink JM, et al (eds) Simulation and Synthesis in Medical Imaging. Springer International Publishing, Cham, pp 154–165, https://doi.org/10.1007/978-3-031-16980-9_15

[150] Kim BN, Dolz J, Jodoin PM, et al (2021) Privacy-Net: An Adversarial Approach for Identity-Obfuscated Segmentation of Medical Images. IEEE Transactions on Medical Imaging 40(7):1737–1749. https://doi.org/10.1109/TMI.2021.3065727, conference Name: IEEE Transactions on Medical Imaging

[151] Bdair T, Saadeh H, Qaqish B, et al (2024) Medical Image-to-Image Translation with Spatial Self-Attention for Radiotherapy in Federated Learning. In: 2024 Fifth International Conference on Intelligent Data Science Technologies and Applications (IDSTA), pp 103–110, https://doi.org/10.1109/IDSTA62194.2024.10746985

[152] Chen S, Jin T, Xia Y, et al (2024) Metadata and Image Features Co-Aware Semi-Supervised Vertical Federated Learning With Attention Mechanism. IEEE Transactions on Vehicular Technology 73(2):2520–2532. https://doi.org/10.1109/TVT.2023.3313593

[153] Selvaraj AK, Prathiba SB, Deepak Kumar A, et al (2024) Co-Training-Based Personalized Federated Learning With Generative Adversarial Networks for Enhanced Mobile Smart Healthcare Diagnosis. IEEE Transactions on Consumer

Electronics 70(3):6131–6139. https://doi.org/10.1109/TCE.2024.3460469

[154] Abdul Sattar Shaikh A, Bhargavi MS, Kumar C P (2023) Weighted aggregation through probability based ranking: An optimized federated learning architecture to classify respiratory diseases. Computer Methods and Programs in Biomedicine 242:107821. https://doi.org/10.1016/j.cmpb.2023.107821

[155] Hu B, Li Y, Zhang Y (2026) M2PFSL: Multi-task and multi-modal IoMT healthcare system via personalized split federated learning. AD HOC NETWORKS 180. https://doi.org/10.1016/j.adhoc.2025.104032

[156] Alam MU, Rahmani R (2023) FedSepsis: A Federated Multi-Modal Deep Learning-Based Internet of Medical Things Application for Early Detection of Sepsis from Electronic Health Records Using Raspberry Pi and Jetson Nano Devices. Sensors 23(2):970. https://doi.org/10.3390/s23020970, number: 2

[157] Rehman A, Xing H, Feng L, et al (2024) FedCSCD-GAN: A secure and collaborative framework for clinical cancer diagnosis via optimized federated learning and GAN. Biomedical Signal Processing and Control 89:105893. https://doi.org/10.1016/j.bspc.2023.105893

[158] Fang ML, Dhami DS, Kersting K (2022) DP-CTGAN: Differentially Private Medical Data Generation Using CTGANs. In: Michalowski M, Abidi SSR, Abidi S (eds) Artificial Intelligence in Medicine. Springer International Publishing, Cham, pp 178–188, https://doi.org/10.1007/978-3-031-09342-5_17

[159] Wang J, Zhang K, Guan Z, et al (2025) Research on the Application of Federated Learning Based on CG-WGAN in Gout Staging Prediction. Computers 14(11). https://doi.org/10.3390/computers14110455

[160] Mamun MMRK, Sherif A, Elsersy M, et al (2025) A distributed deep learning approach for blood sample-based early detection of dementia. IMAGE AND VISION COMPUTING 164. https://doi.org/10.1016/j.imavis.2025.105685

[161] Weldon J, Ward T, Brophy E (2021) Generation of Synthetic Electronic Health Records Using a Federated GAN. https://doi.org/10.48550/arXiv.2109.02543, arXiv:2109.02543 [cs]

[162] Srinivasulu A, Agrawal A, Vedaiyan R (2025) Overcoming diagnostic and data privacy challenges in viral disease detection: an integrated approach using generative AI, vision transformers, explainable AI, and federated learning. FRONTIERS IN VIROLOGY 5. https://doi.org/10.3389/fviro.2025.1625855

[163] Brophy E, De Vos M, Boylan G, et al (2021) Estimation of Continuous Blood Pressure from PPG via a Federated Learning Approach. Sensors 21(18):6311. https://doi.org/10.3390/s21186311, number: 18

[164] Li F, Wang Y, Beyan O, et al (2024) Voxel-Wise Medical Image Generalization for Eliminating Distribution Shift. ACM Trans Knowl Discov Data 18(7):167:1–167:16. https://doi.org/10.1145/3643034

[165] Huang J, Duan T (2025) Deep learning-based approaches for cellular mechanics analysis and secure data sharing in biomechanics. Molecular & Cellular Biomechanics 22(4):1059–Article No.: 1059. https://doi.org/10.62617/mcb1059

[166] Zhao L, Huang J (2023) A distribution information sharing federated learning approach for medical image data. Complex & Intelligent Systems 9(5):5625–5636. https://doi.org/10.1007/s40747-023-01035-1

[167] Lakrouni S, Bah S, Sebgui M (2025) Generated Preserved Adversarial Federated Learning for Enhanced Image Analysis (GPAF). Computers, Materials and

Continua 85(3):5555–5569. https://doi.org/10.32604/cmc.2025.067654

[168] Lv J, Kim BG, Li K (2025) Fed-HealthGen: A Generative Federated Framework for Privacy-Preserving Personalized Healthcare Using Wearable Consumer Electronics. IEEE Transactions on Consumer Electronics 71(4):11440–11452. https://doi.org/10.1109/TCE.2025.3624693

[169] Sikandar M, Din IU, Almogren A (2024) Integrating Generative AI and Federated Learning for Privacy Preserved Sequence-Based Stomach Adenocarcinoma Detection. IEEE Transactions on Consumer Electronics 70(3):5278–5285. https://doi.org/10.1109/TCE.2024.3423786

[170] Apellániz P, Parras J, Zazo S (2025) Improving Synthetic Data Generation Through Federated Learning in Scarce and Heterogeneous Data Scenarios. Big Data and Cognitive Computing 9(2). https://doi.org/10.3390/bdcc9020018

[171] Singhal K, Azizi S, Tu T, et al (2023) Large language models encode clinical knowledge. Nature 620(7972):172–180. https://doi.org/10.1038/s41586-023-06291-2

[172] Luo R, Sun L, Xia Y, et al (2022) BioGPT: generative pre-trained transformer for biomedical text generation and mining. Briefings in Bioinformatics 23(6):bbac409. https://doi.org/10.1093/bib/bbac409

[173] Rombach R, Blattmann A, Lorenz D, et al (2022) High-Resolution Image Synthesis with Latent Diffusion Models. IEEE Computer Society, pp 10674–10685, https://doi.org/10.1109/CVPR52688.2022.01042

[174] Szeliga A (2023) A comparative Study of Deep Generative Models for Image Generation. PhD thesis, Hochschule Hannover, https://doi.org/10.25968/opus-2970

[175] Maslej N, Fattorini L, Perrault R, et al (2024) Artificial Intelligence Index Report 2024. https://doi.org/10.48550/arXiv.2405.19522, arXiv:2405.19522 [cs]

[176] Khan W, Leem S, See KB, et al (2025) A Comprehensive Survey of Foundation Models in Medicine. https://doi.org/10.48550/arXiv.2406.10729, arXiv:2406.10729 [cs]