4 Generative AI for IoT Data Synthesis and Anomaly Detection

Bharathi GR[1*], Ashita Priyadarshini[2] & Madhuri GR[3]
[1&2]Assistant Professor, St Francis College.
[3]Assistant Professor, Kuvempu University.
*Corresponding Author: bharathi.gr7@gmail.com

Abstract

Massive data streams are produced by wearables, sensors, and industrial systems as a result of the Internet of Things' (IoT) explosive growth. However, these datasets frequently encounter issues like noise, imbalance, and scarcity, which restrict the effectiveness of traditional machine learning models. By creating realistic IoT data and facilitating reliable anomaly detection, generative artificial intelligence (Generative AI) provides a workable solution. By learning the underlying data distribution, methods such as diffusion models, variational autoencoders, and generative adversarial networks (GANs) can produce synthetic datasets that maintain privacy and usefulness. This chapter examines how generative AI can be used to solve IoT problems in various fields. It facilitates ongoing monitoring and individualized diagnosis in the medical field. By creating rare fault cases, it facilitates predictive maintenance and defect detection in manufacturing. In energy systems, synthetic data improves resource optimization, while in finance, it improves fraud detection by modelling anomalous patterns. Generative AI improves IoT-driven decision-making and opens the door to a more robust, intelligent, and sustainable digital ecosystem by bridging the gap between data synthesis and anomaly detection.

Keywords: Predictive Maintenance, Generative Artificial Intelligence, IoT Data Synthesis, Anomaly Detection, Synthetic Data.

Introduction

Overview of IoT Data and Challenges

The fast adoption of IoT in the real world has led to the development of a vast volume of different data, known as big data, which is difficult to manage and maintain.

Applications in a range of disciplines, including healthcare, manufacturing, and other industries, as well as energy management, make use of data generated by various connected devices. Because each of these industries must create precise applications, the data generated by IoT devices is interesting. However, there are concerns about IoT-generated data, such as imbalance, shortage, and security. Because of the dynamic nature of the IoT environment and the scarcity of tagged attack data, it is difficult to spot anomalous behaviour and malicious attacks on IoT devices that could jeopardize security.

Role of Generative AI

Generative artificial intelligence (AI) has developed as a formidable paradigm in recent years. It uses massive datasets and clever algorithms to generate new information that is comparable to the original. Generative artificial intelligence (AI) fills gaps in IoT device data caused by challenges such as imbalance and shortage. Gen AI makes data synthesis easier by controlling data shortages, balancing biased datasets, and boosting training datasets via data augmentation. Gen AI helps discover irregularities by learning new patterns from existing data and understanding how the data deviates.

The importance of generative AI for anomaly detection and IoT data synthesis is discussed in this chapter. This chapter initially explores the characteristics of IoT and gives an overview of generative AI models. Next, we look at how they can be applied in a variety of fields. Before outlining prospective areas for additional inquiry, emphasize the use, benefits, and challenges.

Characteristics and Challenges of IoT Data

- **IoT Data Characteristics**

Heterogeneous Data: As the IoT devices use different hardware platforms, the data generated is also in various forms, such as text, images, etc.

  - **Voluminous:** Since IOT uses sensors, a huge amount of data is captured every second, and the data generated adds volume. Generally, the data generated will be in terabytes, petabytes, and zettabytes.
  - **Dynamic and scalable:** IoT devices capture real-time data; hence, they're dynamic in nature and need scalable storage devices due to the increase in data storage. Scalability enables storage for the increasing amount of data generated by IoT devices. It can make use of distributed or cloud storage.
  - **Velocity and Veracity:** The speed at which the data is generated. Since a large amount of data is generated at high speed, the quality and accuracy of the data must be ensured to ensure its trustworthiness for data processing.

- **Simple and Synergetic:** The data generated from IoT is easy to use, deploy, and share and managed by improving the deployment efficiency.
- **Syncretic—Integrated networks:** IoT forges different network types, including ground, aerial (drones), and satellite. It results in hassle-free communication on all networks.
- **Security—Strong protection:** Our conventional security is stationary and slow to catch on to threats. The upcoming IoT security will be active and precise in its proactive forced deployment to counter the attacks that have evolved.
- **Shared:** Current IoT systems often replicate activities across verticals (being developed in the "chimney-like" fashion). IoT enables easy data and resource sharing, eliminating any redundant effort and expense.

- **Key Challenges**

The present IoT landscape is developing quickly; however, it continues to be challenging in many areas. With regard to healthcare, the most important challenges include the accuracy, automation, and reliability of disease detection and remote monitoring, as well as the security and interoperability of patient data. Fog computing is required to compensate for the latency and reliability issues of cloud-only systems, while in 5G networks, traditional RFID tags are costly, less environmentally friendly, and infeasible for long-range operation, imposing the need for greener, cheaper, and more efficient alternatives. Vehicular IoT is subject to trust, privacy, and secure data dissemination, whilst NOCs are exposed to permanent faults that jeopardize the well-functioning of the system, and fault-tolerant designs are necessary. When applied to smart cities, IoT also has to cope with issues for mobility in transport services, sensor faults, vast device networking, and user context security and privacy. Smart agriculture highlights the issues of sustainability, such as water saving, soil quality maintenance, and emission reduction, and data analytics associated with the IoT should focus more on the problem of processing massive, redundant data efficiently, subject to an AI-based solution. Implementation issues, such as difficulties in integrating IT and OT systems, security risks, and a lack of unified data standards for industrial IoT, as well as small device size/low energy consumption vs. high-quality video/VR for multimedia IoT, are also to be solved, and these potentials cannot be explored due to the underutilization of spectrum resources. Overall, there are key challenges in IoT, which include scalability, latency, spectrum sharing, energy consumption, security and privacy, and trust (collectors are also victims of faults).

Generative AI: Foundations

Generative AI integrates with Artificial Intelligence (AI) as part of its base and data within machines to assimilate itself to human abilities of problem-solving,

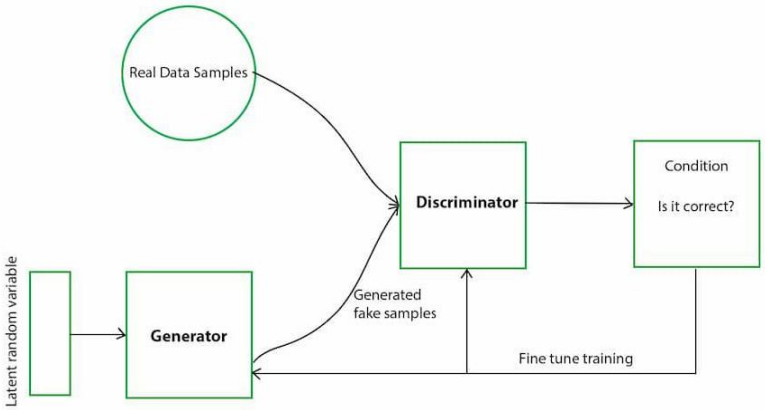
reasoning, and making decisions. AI as a discipline breaks down to Machine Learning (ML), which is a system of training a system using data and making it improve by itself with no direct programming. For instance, supervised ML models work by categorizing emails as spam and not spam, while reinforcement learning trains agent systems like AlphaGo to learn difficult games by repeated attempts. The evolution of ML came with profound deep learning, a concept rooted in artificial neural networks (ANNs) with many layers and central to dominating hard problems like speech and image recognition, as well as self-driving cars. The profound breakthrough in deep learning came with the introduction of the transformer architecture in 2017, which employs attention windows to concentrate on the most pertinent aspects of the data, solving challenges posed by earlier models such as RNNs and LSTMs in managing lengthy sequences. With this, large deep learning models like GPT-3 were developed and trained on large datasets with 175 billion parameters, allowing them to generate texts, codes, and summaries almost as a human would.

Real-world applications of generative AI go beyond data interpretation to include the production of original pieces based on the previously mentioned innovations.

- **Generative Adversarial Networks (GANs)**

A type of deep-learning model first proposed by Ian Goodfellow and his collaborators back in 2014 is known as Generative Adversarial Networks, or GANs for short. With the introduction of GANs, machines can now understand the distribution of the data and synthesize samples on their own. This is unlike the traditional “one and done” models, which only predict or classify. GANs can synthesize images, texts, or audio. They can create new samples that are of very high resemblance to the real world.

A GAN consists of two main components, which in this case are the GAN’s generator and the discriminator. These two components are said to be trained in a competitive manner.



- The generator in this case tries to create real data from a random piece of noise by producing synthetic data.
- The discriminator, meanwhile, is a classifier that attempts to distinguish real data (from the training set) and fake data (the data that the generator synthesizes).

The training, for example, can be considered as a game. In this game, the generator is improving on its “fooling” game, while the discriminator is concentrating on improving their fake sample detection abilities. With the passage of time, this adversarial training is known to produce data that is of very high realism.

- **Variational encoders, or VAEs**

One kind of generative AI model used in deep learning to produce fresh data is called a variational encoder. The original data or a sample of the data that is available is comparable to the data produced by VAE. VAEs consistently produce distinct data according to their prior training. VAEs primarily work with text, audio, and video content.

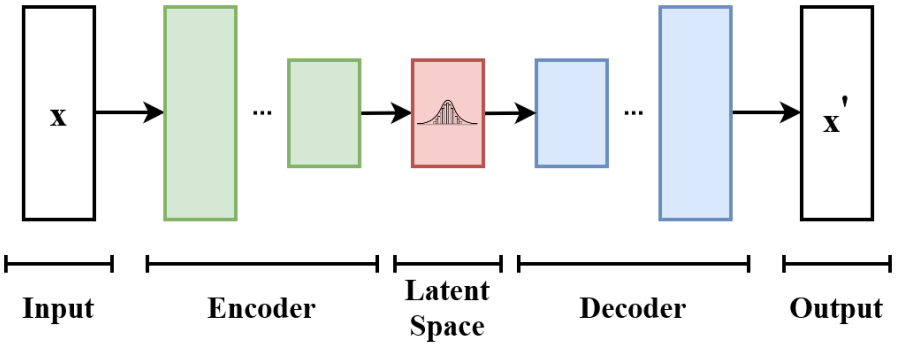
Compressing and decompressing the generated data is essentially the function of the conventional autoencoder. The encoder and decoder are its two components.

In essence, encoders transform the generated data into a representation in latent space.

The encoded data is transformed back into its original form using decoders. VAEs and conventional autoencoders differ slightly.

In order to assist them create new data, VAEs learn from the input data's probability distribution rather than the actual data. The mean (μ) and standard deviation (σ), which characterize the probability distribution in the latent space, or hidden space, are the two forms of data that the encoder produces in VAEs. The original data is regenerated by the decoder using a sample point that is taken from this latent space.

The compressed concealed area where the data resides after encoding is known as latent space. This area is meaningful and continuous. This allows for easy interpolation because comparable input types are clustered and maintained together.



The latent space likewise follows the smooth and clean normal distribution, and training the VAEs entails attempting to reconstruct more precise data production. The loss function provides a clear understanding of VAE performance. Reconstruction loss and KL divergence are the two kinds of loss functions that VAEs employ. How closely the generated output resembles the original input is determined by reconstruction loss. Resemblance between the newly generated data and the old data which may aid in improving the decoder's training.

As the data sample point is gathered from the latent space representation, the KL Divergence loss function calculates the degree to which the learnt distribution deviates from the normal distribution. It promotes the uniform, smooth latent space. Reconstruction Loss + KL Diversion Loss = Total Loss

The output of VAEs is dispersed; random sampling is used. Each time, the created data is unique.

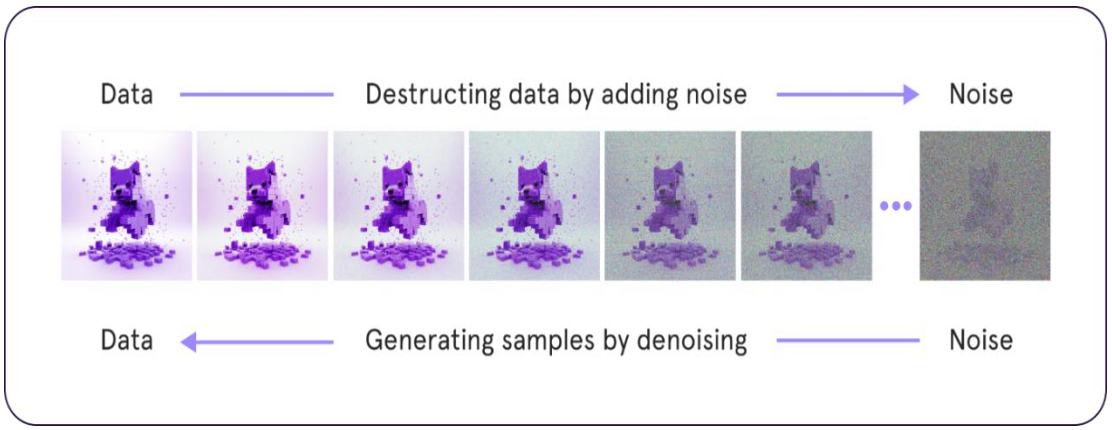
- **Uses for VAEs**
  - It can be used to create clothes patterns in the fashion business.
  - It can be applied to data augmentation in the medical profession.

- **Diffusion Model**

The generative AI models known as diffusion models, which produce data through progressive learning, are widely utilized to produce text-to-images that are realistic, varied, and of high quality.

They operate in two distinct stages:

- **Forward Diffusion:** In this model, noise is continuously introduced to a clean image over brief periods of time, resulting in a corrupted or entirely noisy image.
- **Backward Diffusion:** In this model, a full noisy image is captured, and the noise is gradually eliminated to produce a new, clear image.



The diffusion probability density model is a probabilistic process that forms the basis of diffusion models. At a specific time, point in the diffusion process, this model indicates the likelihood that a clear image will transform into a noisy image version.

Some features were more significant when the photos were denoised, so researchers used this information to create a technique called Refusion, which allows them to choose the optimal noise level to train a smaller, faster model that can aid in other image classification and segmentation procedures. The performance of other AI models is also improved by this diffusion model. demonstrates that diffusion models are excellent for learning significant visual aspects in addition to producing images.

Transformers and UNET are two different architectures that are used for this learning. The first step of the process is gathering the bid data, which is varied and can be used to identify underlying trends and produce high-quality patterns. To what extent are vast quantities of data made available for training? The model will produce so many correct outcomes. A diffusion model allows us to produce two different types of images.

- **Unconditional Images**: In this model, the noise is converted straight into any random image without any input.
- **Conditional Pictures:** In this case, the model is given additional data, such as class labels or text descriptions, that might direct the model to produce particular kinds of images.

While this methodology is effective at producing text-to-image sample data, it is a little more difficult to produce data from audio and video.

Generative AI for IoT Data Synthesis

- **Synthetic Data Generation:** Data that is purposefully produced to resemble the data in the original datasets is known as synthetic data. Due to privacy and security concerns, as well as the limited amount of data in the original dataset, the synthetic data was necessary. In these circumstances, synthetic data is

essential to the application, which needs a large amount of data to demonstrate improved performance, in contrast to the actual dataset. The many generative AI models produce synthetic data.

For instance, sensors in the manufacturing sector provide data on temperature, vibration signatures, and production parameters that are useful for operational optimization research and predictive maintenance.

- **Synthetic Data Types**
  - **Tabular Information: This** type of dataset falls under the category of structured data, which is frequently utilized in database and Excel applications, and is saved as rows and columns.
  - **Time Series Data:** This type of data is similar to the original dataset because it is captured throughout time. The relevant sensors record the data in a sequential fashion. Artificial intelligence is trained using this type of synthetic data to learn how to replicate patterns in actual data. ECG data, for instance, can be used to identify cardiac issues.
  - **Image Data:** Artificial still images produced for computer vision methods such as surveillance systems, medical imaging, and object detection.

  An example would be the use of an Internet of Things camera for traffic monitoring, which facilitates effective traffic analysis.
  - **Textual Data:** To aid in NLP and IOT log analysis, synthetic text data that is comparable to the real content is developed.

  Example: Notification of an alert in the event that a machine detects motion of any kind
  - **Network and graph data:** For example, several IoT devices connected to one another in a network in a smart factory are examples of synthetic data that depicts the network connection or relationship used in various platforms. This aids in the industry's network optimization.
  - **Hierarchical/Semi-structured data:** This type of data is sophisticated and nested, and it is saved in XML/JSON format.

  For instance, a smart washing machine's JSON file can be used to model intricate IoT systems.
  - **Video Data:** A collection of artificial image frames produced over time for tasks involving the creation and analysis of videos.

  For instance, a surveillance system that records the number of individuals who enter a site and their activities over a specific time period aids the AI model in detecting anomalies in the recorded video frames.

- **Methods Employed**
  - Traditional and Rule-Based Approaches
  - Conventional machine learning techniques like the LSTM neural network and the Markov model
  - Deep learning techniques like AAEs, VAEs, and GANs
  - Big language models like BERT, GPT-1, and GPT-2, among others.
  - IOT-specific frameworks and tools like the Synthetic Data Vault and Great AI.

- **Data Augmentation**

By including actual types of data produced by generative AI models, data augmentation involves adding additional information to the dataset's existing data in order to boost its availability.

By training the models to perform better with dependable accuracy, data augmentation lets us get over the privacy and scarcity issues with the actual data from IoT devices.

Data augmentation is necessary to address issues with data generated by IoT devices, including limited or unbalanced sensor data, IoT device failure at particular times, and user data privacy concerns.

- **IoT Data Generation Using Generative AI**
  - **Gathering and Preparing Data:** Gathering unprocessed data from sensors or Internet of Things devices, cleaning it up, turning it into a time series, or sensor logs, and then formatting it according to the model being used.
  - **Model Selection:** A suitable generative AI model is chosen based on the type of data and its requirements. For realistic data creation, a GAN is used, whereas VAEs are used for probabilistic sampling and latent space search.
  - **Developing a Model for generative AI:** Using adversarial training, the AI models are trained on the real data that is accessible, taking into account the necessary privacy considerations, in order to produce new data that is comparable to the original data.
  - **Synthetic Data Generation:** The process of creating new samples that are comparable to the original data while ensuring that they are all real-type data and transforming them into the original data's format is known as synthetic data generation.

- **Data Integration and Augmentation**

To enhance the training dataset, the generated samples are combined with actual data. By offering a vast amount of data to obtain precise conclusions and addressing the problems of data scarcity, it contributes to the enrichment of the original data set.

- **Model Training:** AI models for anomaly detection and image recognition applications are trained using the enriched dataset. where the model continuously improves its accuracy, precision, recall, and F1 score by learning new patterns.
- **Privacy-preserving data sharing:** Data sharing that protects privacy: Generative AI modes such as GAN and VAEs create synthetic IOT data because the data produced by IOT devices may be restricted, sensitive, or subject to legal concerns. This way, the models' synthetic data can be used to train the models without jeopardizing the privacy concerns of the companies.

Consider using ECG signals to track a patient's heart-related issues. Generative AI models produce comparable data that can be used to identify heart-related issues without disclosing the patient's actual information while maintaining privacy concerns. While users want their information to be private, businesses need data to better their goods and services.

Some privacy-preserving algorithms have been created to address these kinds of problems since, in certain cases, an AI model may, after being trained, output data that is an actual copy of real data, thereby violating privacy.

- **Privacy-Preserving Techniques**
  - **Differential Privacy (DP):** Techniques for protecting privacy include: 1. Differential Privacy (DP): This technique involves adding noise to the original data in order to conceal the true information because of privacy concerns about personal information. The goal of differential privacy is to create data that is comparable to the original data without altering it in any way. This is accomplished by introducing noise or randomness into the data, which creates new data. Unlike earlier approaches that used grouping and masking the data, no one can determine if this information is a part of the original data or not. Differential privacy offers an organized mathematical approach to guarantee the original data's privacy protection. More noise will result in less accurate data, but privacy will be protected. Conversely, less noise will result in more accurate results, but privacy will be compromised. Therefore, in order to achieve the strong privacy

approach, we must regulate and balance the appropriate quantity of accuracy and noise.

- **Federated Learning:** Federated learning is a machine learning approach that, in order to preserve privacy concerns, trains the model locally using datasets without sharing them on a central server. When federated learning is used, the model is trained locally rather than using data. This is known as a global model, and other participants train the model using their own local data, such as hospital, bank, or IoT models, among others. Each participant trains the global model without sharing their data; only knowledge is transferred. Because personal information never leaves the area, this technique lowers the possibility of privacy invasion. Differential privacy is typically combined with federated learning. Secure aggregation should prevent data breaches in order to produce a better model.
- **Homomorphic Encryption (HE):** One kind of cryptography model that enables you to perform certain computations on your encrypted data without even knowing the data itself is homomorphic encryption. When your code is decrypted, it will be identical to the calculation made on the original data. In this way, homomorphic encryption helps businesses with their privacy concerns by protecting their data without disclosing important information.

  In remote, dynamic, and untrusted environments—like cloud environments, industry IoT, etc.—homomorphic encryption is typically employed.

For instance, depending on their needs, a healthcare company may submit patient data to a cloud provider for statistical analysis. The data is transmitted in an encrypted manner, and the cloud service does the analysis without ever viewing the data, yielding precise findings. Homomorphic encryption can be used to accomplish this.

For instance, depending on their needs, a healthcare company may submit patient data to a cloud provider for statistical analysis. The data is transmitted in an encrypted manner, and the cloud service does the analysis without ever viewing the data, yielding precise findings. Homomorphic encryption can be used to accomplish this.

- **Secure Multi-Party Computation (SMPC):** Secure Multi-Party Computation is a cryptographic technique that allows numerous parties to collaborate to compute the combined results without disclosing their actual data to one another. This is accomplished by encrypting and sharing one's data with other parties, which prevents others from identifying one's actual data while enabling collaboration on the final, accurate results.

When working on projects that are dispersed across several organizations, SMPC is quite helpful.

For instance, a number of hospitals wish to forecast a specific disease diagnosis without disclosing patient information, but by integrating all patient data to verify the disease analysis, each hospital will use SMPC to exchange encrypted protected values with one another in order to obtain an average disease prediction that does not violate privacy laws.

Secret sharing, corrupted circuits, and homomorphic encryption are common SMPC approaches.

Generative AI for Anomaly Detection

- **Pipeline for Anomaly Detection**

Finding outliers is the essence of anomaly detection. The generative AI models are responsible for these detections. These models detect irregularities in intricate settings like the cloud, IoT, and industries.

The procedures for carrying out anomaly detection area**.**

- **Collect and prepare data streams:** Gather a lot of data from various environments, clean it up by identifying missing values, normalize it, and separate it properly to make sure it works with the model architecture. Then, break the data into fixed stream groups to find patterns.
- **Data Augmentation:** If there aren't many anomalies in the original data set, we can use a generative AI model called GAN to produce new abnormalities. We can find the anomalous pattern in the original data set by using this model to create new anomalies. For instance, we can observe data probing in IoT devices to detect data attacks. By adding the synthetic data, data augmentation enables us to spot the less suitable patterns in the original data.
- **Model Training:** The GAN model of generative AI has been trained. A discriminator and a generator are the two parts of the model. The discriminator's task is to determine if the data produced by the generator is authentic or fraudulent, whereas the generator's task is to create the data samples at random using synthetic data. It is a generator loss if the discriminator detects it as phony data. With the aid of this model, the discriminator and generator compete with one another to enhance the module and produce better synthetic data if the discriminator recognizes it as real data. We can anticipate some sort of discrepancy if the generated data differs significantly from the actual data.
- **Feature representation and inference:** The model is prepared to recognize the data as normal data when it has been trained. The model will

look for any abnormalities when each new data point is run through it. An anomaly is recognized if the data is out of the ordinary. Take the number seven as an example of a handwritten number. The model has been trained with thousands of handwritten images. After training, a fresh image is provided as 3, but the model attempts to replicate it as 7, which does not resemble the image of 7, thus producing an anomaly.

- **Anomaly score:** calculating an anomaly score for every data instance produced based on standards such
  - Reconstruction loss measures the model's ability to reconstruct the input data. The anomaly will be revealed by how dissimilar the generated data is from the original data. The anomaly increases with the size of the divergence between the created and original data.
  - **Discriminator loss:** this indicates the likelihood that the data produced is derived from the actual distribution; the less likely this is, the more anomaly occurs.
  - **Statistical comparisons:** the latent vector distances and the indexes have similarities.

To determine whether the model can detect the anomaly or not, a threshold anomaly score must be determined.

- **Deployment in the real-time system:** In order to manage idea drift and changing behaviour patterns, use the trend anomaly detection framework in a real-time setting, such a cloud platform, where incoming data streams are continuously observed and assessed for anomalies. As real-world data changes, Rustam may integrate techniques like sliding window analysis, allowing for prompt and precise detection.
- **Alert and improve:** The system sends out an alert for the security team to look into or to start automated mitigation procedures when the anomaly score exceeds the predetermined level. Over time, the generative model's accuracy and resilience can be improved by using feedback from confirmed occurrences, whether they are actual attacks or false positives.

- **Detection Techniques**

In the Industrial Internet of Things (Iao), malfunctions of device and sensor streams and cyberattacks make security and dependability critical, hence the importance of anomaly detection. IoT data, unlike most information, is so disproportionate that traditional methods function badly. Developments in generative AI have come up with effective solutions.

Applies generative adversarial networks (GANs) to train with the normal IoT traffic and learn the distribution of plausible behaviour. During inference, the GAN

marks the input as an anomaly if it cannot reconstruct it with sufficient accuracy. For time-series data, Tad GAN and similar models can reconstruct normal sequences of sensor readings, with deviations in reconstruction error signifying an anomaly — powerful methods for fault detection and predictive maintenance.

GANs are invaluable for balancing datasets of different distribution (augmentation) for uncommon attack types, such as probing and misconfigurations. When GAN-based augmentation is applied with deep reinforcement learning (DRL) powered intrusion detection systems, the detection accuracy of minority attacks is significantly enhanced. Increased robustness is achieved with hybrid approaches combining GANs and other autoencoders, reinforcement learning, as well as other GAN augmentations

- **Generative AI for Anomaly Detection**

  Multiple applications in this field include:

  - **Healthcare:** Utilizing medical images, generative models like GAN, VAEs, and diffusion models detect patterns for diagnosing diseases, augment rare lesions, and localize anomalies.
  - **Cyber security:** Detecting unusual data within logs and time series, as well as identifying malware and network faults.
  - **Manufacturing:** Recognizing machine and product defects while incorporating synthetic data to enhance model training.
  - **Finance:** Identifying fraudulent transactions and detecting abnormal activities in insurance and loan processes.
  - **IoT/Smart Cities:** Detecting unusual activities or sensor readings in video surveillance systems to identify anomalies.

Domain-Specific Applications IOT

Electronic Health Systems

The healthcare industry demonstrates a leading example of successful IoT integration for its daily practices. Medical services use IoT primarily to retrieve information at high speeds. The medical field maintains a continuous development process that focuses on human health preservation through disease identification and prevention as well as treatment management. The medical field extends its operations outside hospitals through equipment delivery and insurance document administration. The Internet of Medical Things (IoMT) emerged from medical IoT advancements which now provides expanded potential for enhanced and accessible healthcare services. IoMT connects medical devices to IoT components that enable detection alongside computational capabilities. The medical field employs IoT technology through multiple applications which include integrated situational forecasting and integrated portal setup and remote urgent care services and intelligent medical entry

and wearable access and reactions to harmful medications. Implantable and wearable sensors including electronic capsules and intelligent nutrition trackers and automated bedding and assistive tools serve as primary drivers of IoT progress in medical applications. Small devices which monitor health through heart rate and blood alcohol content tracking and hypertension detection and body temperature measurement represent key elements of medical IoT development. Modern technology allows easy access to electrocardiograms and brain wave patterns as well as muscle activity measurements [59]. The process of evaluating injuries became easier due to the implementation of IoT-injury assessment platforms. The Internet of Things enables smartphone biosensors to detect multiple pathogens which include the detection of SARS-CoV-2. Through the combination of these innovations' healthcare professionals including physicians and caregivers can generate urgent medical decisions and decrease costs by accessing multiple data sources through immediate online connections. Medical assistance can be provided instantly through IoT devices that accept voice commands. Medical professionals can now access patient medical histories and medication records which contain details of illness progression and medication compliance. Pharmaceuticals use barcode labelling systems to enable direct patient delivery. Emergency vehicles connect through satellite navigation (GPS) systems to radio-frequency identification (RFID) systems for accelerated patient response times.

The health data integration through IoT tools has established multiple operational challenges. The primary challenges exist in constant data accessibility, large data volume maintenance and storage, power usage, asset compatibility issues, privacy protection measures and data anonymization, and universal system availability. Cloud-based solutions provide solutions to numerous issues but they create increased energy consumption. Different types of IoT devices create data structure inconsistencies throughout IoT systems. Field personnel become vulnerable to risks when electronic health lacks sufficient collaboration since they must handle multiple separate devices that use different software systems. Superior medical care cannot depend solely on IoT tools when there is a deficiency of healthcare professionals who possess the necessary expertise to use these tools effectively the successful operation of these devices along with their output analysis depends on proper handling and analysis by users. The medical use of IoT enjoys broad public adoption but information management remains mostly with skilled users while older patients require guidance to use modern IoMT tools for self-monitoring. The healthcare sector has the potential to save substantial funds by properly implementing IoT systems. To achieve maximum benefits from IoT implementations all participants need to follow a targeted strategic approach while conducting extensive research and providing proper instructions to medical staff as well as achieving standardization across various IoT devices.

Healthcare institutions that advance technology infrastructure create essential new approaches for disease prevention and management and improve their adaptability during public health crises. Healthcare systems will achieve enhanced operational efficiency together with increased capabilities through these innovations which will produce superior public health outcomes. Among the most encouraging developments in healthcare technology is the Internet of Medical Things (IoMT).

The Internet of Medical Things (IoMT) represents an interconnected framework of devices and systems aimed at gathering, analysing, and distributing health-related information. Medical services now operate differently because this innovation enables remote patient monitoring and diagnostic and therapeutic procedures. IoMT takes the core principles of Internet of Things (IoT) technology and modifies them specifically for use in healthcare environments to meet their unique needs and challenges-through the application of Medical Internet of Things (IoMT) healthcare organizations can build connected data systems which link smart devices including hospital devices diagnostic equipment and wearable sensors. The core elements of IoMT continuously gather and generate health information which central hubs process for detailed analysis. The obtained knowledge functions as an essential support tool for medical professionals when making their decisions.

During recent years the IoMT has gained widespread usage across various healthcare domains which include illness detection and patient monitoring from afar and smart clinics and epidemic surveillance. The wide adoption demonstrates how IoMT can transform healthcare delivery systems and improve patient outcomes.

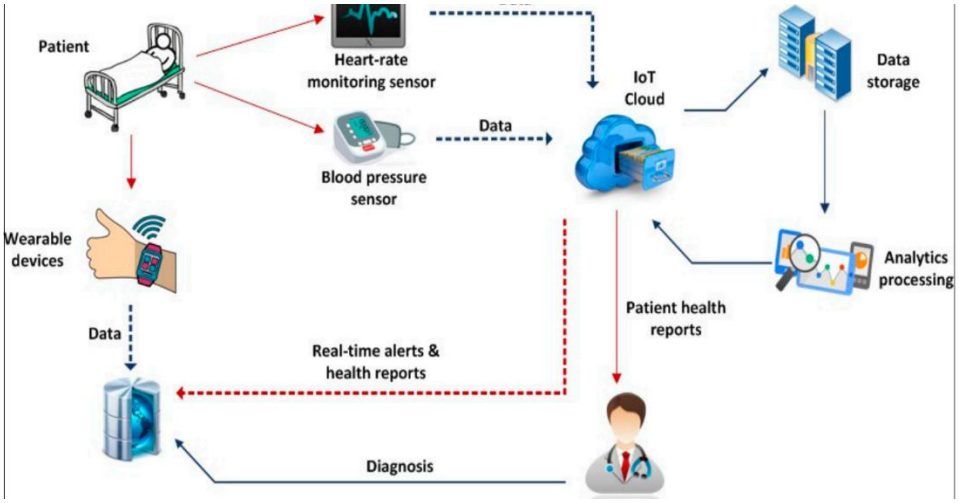
The current research demonstrates limited investigation into the precise ways IoMT addresses medical field challenges despite growing interest in its adoption. Existing research studies about IoMT provide general overviews about its applications yet they fail to evaluate how well it meets specific healthcare requirements. The current study addresses this gap by examining the following fundamental elements: Current healthcare systems contain specific challenges which IoMT aims to solve through resolving security issues along with energy optimization and intelligent sensor operation as well as reliable equipment maintenance.

The distinctive features of the IoMT enable successful resolution of healthcare challenges through its capabilities of continuous patient monitoring and real-time data management alongside artificial intelligence and machine learning integration.

The real-world deployment of IoMT demonstrates its practical impact through specific examples which highlight its effects in medical settings. The study examines the implementation of IoMT through detailed assessments of its applications in remote patient monitoring and advanced hospitals and epidemic tracking.

This Chapter seeks to provide a comprehensive understanding of IoMT through expanded elements to fulfil the current research gap about healthcare

applications. Through detailed analysis together with real-world instances we demonstrate how IoMT can transform medical service delivery and improve patient outcomes.



E-healthcare Scenario

The implementation of IoT devices for health information access has generated certain obstacles. The challenges include ubiquitous data access and physical storage of large data volumes and data availability and maintenance and energy consumption and resource interoperability and privacy and security and data anonymity and unified and universal access.

- **Big Data in Healthcare**: Refers to the massive volume of health-related data collected by IoT devices and Remote Health Monitoring (RHM) networks.
- **Computational Demands**: Managing this data requires high-performance computing and extensive storage capacity.
- **Cloud Solutions**: Cloud computing and storage are essential for handling and processing healthcare data efficiently.
- **Patient Confidentiality**: Despite the benefits of cloud systems, protecting patient privacy remains a top priority.
- **Security Challenges**: RHM systems face significant risks related to:
  - Computer and network security
  - Storage and physical security
  - Authentication protocols
- **Data Protection Techniques**:
  - Genetic algorithms
  - Encryption and decryption methods

- **Trust Issues**: Many current security frameworks rely on third-party services, which may not always be reliable or secure.

Healthcare IoT refers to a network of interconnected devices that monitor and transmit real-time health data. These include:

- **Wearables**: Smartwatches, fitness trackers, ECG monitors
- **Smart Medical Devices**: Glucose monitors, insulin pumps, portable ultrasound machines
- **Connected Infrastructure**: Smart beds, remote patient monitoring systems, hospital asset trackers

These devices collect continuous streams of data such as:

- Heart rate, blood pressure, glucose levels
- Body temperature, oxygen saturation
- Sleep patterns, physical activity

This data is rich, but raw — and that's where Generative AI steps in.

Generative AI doesn't just analyse data — it **creates, predicts, and personalizes**. Here's how it enhances Healthcare IoT:

Synthetic Data Generation

- Creates realistic patient data to train models when real data is scarce or sensitive
- Helps simulate rare conditions for research and testing

Anomaly Detection

- Learns normal patterns from IoT data and flags deviations (e.g., sudden drop in oxygen levels)
- Enables early detection of complications like arrhythmias or diabetic emergencies

Personalized Insights

- Tailors' health recommendations based on individual patterns
- Generates adaptive care plans and alerts for patients and providers

Medical Imaging & Diagnostics

- Enhances scans and images from IoT-connected devices
- Generates high-resolution reconstructions or predicts disease progression

Virtual Health Assistants

- Uses patient data to generate conversational support, reminders, and education
- Improves adherence to treatment and boosts engagement

Manufacturing

IoT in Manufacturing/Industrial IoT (IIoT):

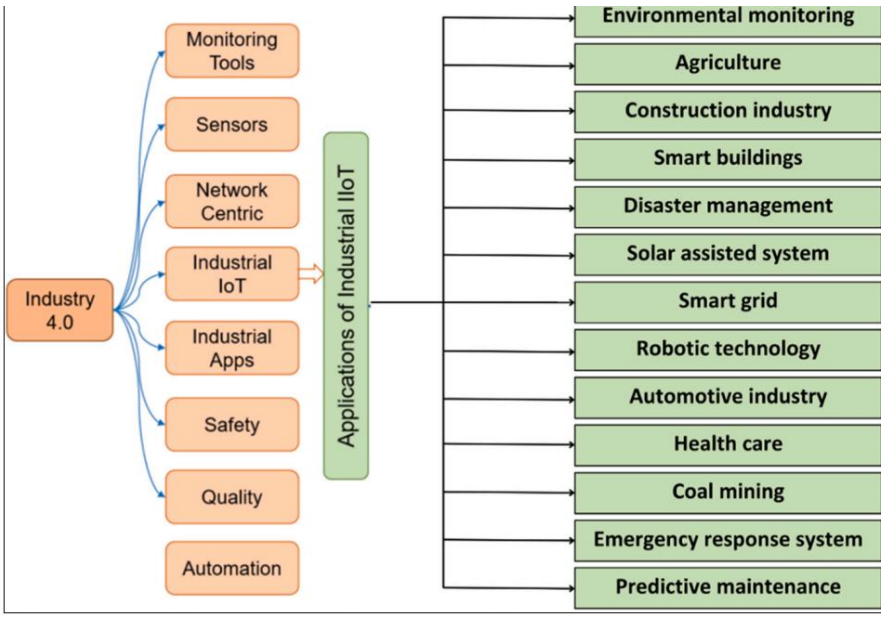
IoT in Manufacturing (also called **Industrial IoT – IIoT**) means connecting machines, tools, sensors, and systems in factories to the internet for **real-time monitoring, automation, and data-driven decision-making**.

Key Applications

- **Predictive Maintenance** – sensors on machines detect vibrations, temperature, pressure → predicting breakdowns before they happen.
- **Smart Quality Control** – IoT cameras/sensors check products during production.
- **Supply Chain Optimization** – track raw materials, shipments, and inventory in real time.
- **Energy Efficiency** – monitor energy usage to reduce costs.
- **Worker Safety** – wearables detect unsafe conditions in factories.

Manufacturing IoT (or IIoT) connects the physical and digital worlds through:

- **Sensors**: Track temperature, vibration, pressure, humidity, etc.
- **Connected Machines**: CNC machines, conveyors, robotic arms
- **Monitoring Systems**: Real-time dashboards for production, energy usage, and safety
- **Robotics & Automation**: Perform precision tasks and adapt to changing conditions.



Generative AI transforms IoT data into **actionable intelligence** by:

Predictive Maintenance

- Learns from sensor data to forecast equipment failures before they happen
- Reduces downtime and extends machine life

Process Optimization

- Simulates production scenarios to find the most efficient workflows
- Suggests adjustments in real time to improve throughput and reduce waste

Anomaly Detection & Diagnostics

- Flags unusual patterns in machine behaviour or energy usage
- Generates diagnostic reports and troubleshooting guides

Synthetic Data for Simulation

- Creates realistic data to test new systems or train models without disrupting operations

Human-Machine Collaboration

- Acts as a virtual assistant for engineers, interpreting dashboards and suggesting actions

Finance

The digital transformation of finance with precision. Let's elevate your summary into a full-fledged view of how **Banking & Finance IoT + Generative AI** is reshaping the financial ecosystem:

Banking & Finance IoT: The Connected Infrastructure

IoT in finance creates a **web of smart, data-generating endpoints**, including:

- **ATMs & POS Machines**: Track transactions, usage patterns, and fraud signals
- **Biometric Sensors**: Enable secure authentication via fingerprint, facial recognition, or iris scans
- **Smart Payment Systems**: Tap-to-pay, QR-based wallets, NFC-enabled devices
- **Mobile Apps & Wearables**: Monitor spending, savings, and financial behaviour
- **Connected Branches**: Share data across locations for seamless customer experience

These systems generate **real-time data** across three key dimensions:

- **Financial**: Transactions, balances, credit scores
- **Behavioural**: Spending habits, location-based activity, device usage
- **Operational**: Queue lengths, machine health, service demand

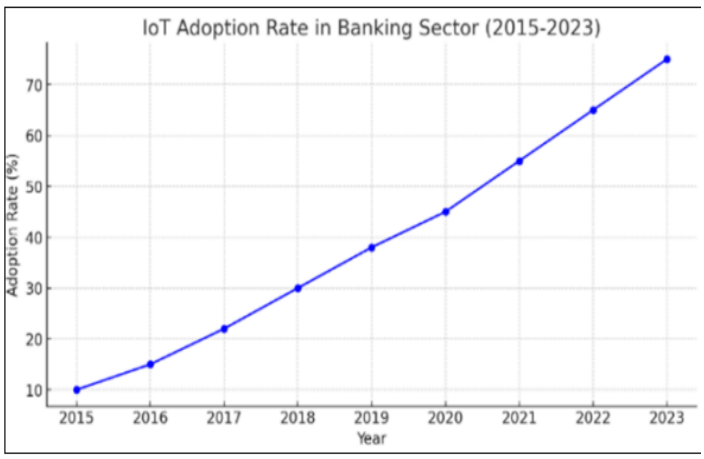


Fig 1: IoT Adoption Rate in Banking Sector (2015-2024)

Generative AI: The Intelligence Engine

Generative AI supercharges this data with **insight, foresight, and personalization**:

Fraud Detection & Risk Management

- Learns from transaction patterns to flag anomalies instantly
- Generates synthetic fraud scenarios to test system resilience

Personalized Financial Advice

- Creates tailored savings plans, investment suggestions, and budgeting tips
- Powers virtual financial advisors that adapt to user behaviour

Customer Engagement

- Crafts personalized messages, offers, and nudges based on real-time data
- Enhances chatbots with natural, human-like responses

Security & Compliance

- Automates regulatory reporting and audit trails
- Uses biometric and behavioural data to strengthen identity verification

Operational Optimization

- Predicts ATM cash-out times, POS failures, or branch congestion
- Suggests staffing or inventory adjustments based on usage trends

Real-World Impact

Generative AI could unlock **$200–340 billion annually** in banking productivity gains. Indian banks like RBL and ICICI are already piloting GenAI for customer service, fraud detection, and loan processing.

Energy and Smart Grids

Smart Grids are modern energy networks that use **IoT sensors, smart meters, and connected devices** to monitor and manage electricity flow.

They allow **real-time energy tracking, demand prediction, outage detection, and efficient distribution**.

These IoT systems form the backbone of **smart grids**, which are digital upgrades to traditional power networks. They enable:

- **Real-time visibility** into energy flow and demand
- **Decentralized energy management**, including renewables and microgrids
- **Predictive maintenance** to prevent outages and equipment failures
- **Consumer empowerment** through usage insights and cost control

Smart grids are expected to surpass **$130 billion globally by 2028**, driven by IoT-enabled efficiency and sustainability.

The synergy between **Generative AI and IoT in energy systems** with clarity and precision, Let's elevate it with a few real-world insights and industry signals to show how this is already unfolding globally:

GenAI + IoT in Energy & Smart Grids: Real-World Impact

Your seven-point framework is spot-on. Here's how leading organizations are applying these concepts:

Synthetic Data for Grid Simulation

- The U.S. National Renewable Energy Lab (NREL) is using GenAI to simulate rare grid faults and outages, helping utilities build more resilient systems.
- These synthetic datasets are crucial for training models where historical data is sparse or sensitive.

Demand Forecasting & Load Balancing

- Companies like Hexaware have improved electricity consumption forecasting by **50%** using GenAI models trained on IoT and weather data.
- This helps prevent blackouts and optimize energy distribution during peak hours.

Anomaly Detection

- GenAI is being used to detect cyber threats and power theft by learning normal grid behaviour and flagging deviations in real time.
- This ensures both **security and operational continuity**.

Renewable Integration

- AWS and Total Energies are using GenAI to simulate solar and wind output, enabling smarter decisions about when to use stored vs generated energy.
- This is key for balancing intermittent renewables with grid stability.

Digital Twin Technology

- Utilities are building GenAI-powered digital twins of their grids to simulate failures, surges, and cyberattacks before they happen.
- These virtual models improve planning and reduce downtime.

Energy Trading & Pricing

GenAI is helping simulate dynamic pricing models for peer-to-peer energy trading, especially in smart communities with solar rooftops.

- This promotes fair pricing and efficient energy use.

Predictive Maintenance

- IoT sensors on turbines and transformers feed data into GenAI models that predict failures, reducing repair costs and improving uptime.

This fusion of GenAI and IoT is not just futuristic — it's already reshaping how energy is produced, distributed, and consumed.

Core Components of Smart Grid Architecture

Here's a structured view of the main elements that transform a traditional grid into a Smart Grid:

Component	Function
Central & Decentralized Stations	Manage power generation and distribution across regions
Renewable & Nonrenewable Plants	Feed energy into the grid from solar, wind, coal, hydro, etc.
Grid Automation Infrastructure	Enables real-time control, fault detection, and self-healing capabilities
Intelligent Substations	Monitor and regulate voltage, load, and power quality
Smart Switches & Distribution Automation	Route electricity efficiently and respond to outages
EV Charging Stations	IoT-enabled for load balancing and energy trading
Energy Storage Facilities	Store surplus energy for peak demand or backup

Supporting Technologies

- **Advanced Metering Infrastructure (AMI)**: Smart meters for real-time usage tracking
- **Communication Networks**: Wireless (Zigbee, LTE), fiber optics, and satellite links
- **Data Analytics & AI**: For forecasting, anomaly detection, and optimization
- **Cybersecurity Systems**: Protect grid data and infrastructure from threats

Challenges and Limitations OF Generative AI for IoT Data Synthesis and Anomaly Detection

The core challenges of applying Generative AI to IOT anomaly detection. To build on your summary and help you take this further, here are a few **strategic directions and solutions** that researchers and engineers are exploring to tackle these issues:

Potential Solutions & Research Directions

- **Improving Data Quality**
  - **Data Augmentation**: Use domain-specific transformations to enrich training data.
  - **Self-supervised Learning**: Leverage unlabelled IoT data to learn robust representations.
  - **Federated Learning**: Train models across decentralized devices without sharing raw data.
- **Reducing Computational Load**
  - **Model Compression**: Techniques like pruning, quantization, and knowledge distillation help deploy GenAI on edge devices.
  - **TinyML**: Emerging field focused on running ML models on ultra-low-power hardware.
- **Mitigating Overfitting & Mode Collapse**
  - **Regularization Techniques**: Dropout, weight decay, and early stopping.
  - **Improved GAN Architectures**: Like Style GAN, WGAN-GP, and Unrolled GANs to reduce mode collapse.
- **Better Anomaly Definitions**
  - **Hybrid Models**: Combine rule-based systems with GenAI to encode domain knowledge.
  - **Human-in-the-Loop**: Incorporate expert feedback to refine anomaly boundaries.

- **Enhancing Explain ability**
  - **XAI (Explainable AI)**: Use techniques like SHAP, LIME, or attention maps to interpret GenAI outputs.
  - **Surrogate Models**: Train interpretable models to approximate GenAI behaviour.
- **Securing Synthetic Data**
  - **Differential Privacy**: Add noise to training data to prevent leakage.
  - **Robust Training**: Use adversarial training to defend against attacks.
- **Scaling Efficiently**
  - **Stream Processing Frameworks**: Apache Flink, Kafka Streams for real-time data handling.
  - **Edge-Cloud Collaboration**: Offload heavy computation to cloud while keeping latency-sensitive tasks on edge.
- **Improving Evaluation**
  - **Custom Metrics**: Tailor metrics like FID, precision-recall, and anomaly score distributions for IoT.
  - **Benchmark Datasets**: Push for open, diverse IoT datasets across domains.
- **Boosting Domain Adaptability**
  - **Transfer Learning**: Fine-tune models across domains with minimal data.
  - **Meta-Learning**: Train models to adapt quickly to new tasks with few examples.

Future Research Directions

Data-Centric Modelling

- **Challenge**: Multimodal, irregular, noisy IoT data.
- **Solutions**:
  - Transformer-diffusion hybrids for long-range dependencies.
  - Graph-based generative flows for spatial-temporal sensor modelling.
  - Neural ODEs for irregular time sampling.
- **Benchmarks**: SWaT, WADI, TON_IoT, MSL, NAB.

Detection-Driven Generation

- **Challenge**: Scarcity of labelled anomalies and poor generalization.
- **Solutions**:
  - Unified generative-detection architectures.

  - Conditional anomaly synthesis for adversarial robustness.
  - Contrastive and forecasting-based self-supervision.
- **Metrics**: AUC-PR, latency, false alarms, calibration.

Adaptation & Lifelong Learning

- **Challenge**: Concept drift, few-shot anomalies, evolving environments.
- **Solutions**:
  - Generative replay for continual learning.
  - Meta-learning for rapid domain adaptation.
  - Prompt-based anomaly synthesis.
- **Evaluation**: Forgetting rate, adaptation speed, online drift resilience.

Privacy & Federated Modelling

- **Challenge**: Sensitive, distributed IoT data.
- **Solutions**:
  - Federated GANs and DP-diffusion models.
  - Privacy-preserving training via DP-SGD.
  - Leakage audits and membership inference tests.
- **Metrics**: Privacy risk, utility trade off, communication cost.

Edge Deployment & Efficiency

- **Challenge**: Real-time constraints, limited compute.
- **Solutions**:
  - Model compression (distillation, pruning).
  - Streaming generative models.
  - Hardware-aware design (TinyML, Edge TPU).
- **Metrics**: Latency, energy, memory footprint, detection accuracy.

Security & Robustness

- **Challenge**: Adversarial attacks, poisoning, evasion.
- **Solutions**:
  - Distribution ally robust optimization.
  - Adversarial testing frameworks.
  - Certified anomaly detection bounds.
- **Evaluation**: Attack success rate, robustness under contamination.

Causal & Counterfactual Reasoning

- **Challenge**: Root-cause analysis and actionable insights.

- **Solutions**:
  - Causal generative models with intervention capabilities.
  - Counterfactual simulation for diagnostics.
  - Causal discovery from observational streams.
- **Metrics**: Counterfactual fidelity, diagnostic accuracy, operator feedback.

Human-Centred explain ability

- **Challenge**: Trust, interpretability, operational integration.
- **Solutions**:
  - Sensor-level attribution and temporal saliency.
  - Active learning with synthetic queries.
  - Interfaces for corrective feedback loops.
- **Evaluation**: Trust scores, label-effort reduction, resolution time.

Standardization & Reproducibility

- **Challenge**: Fragmented benchmarks and evaluation protocols.
- **Solutions**:
  - Multimodal benchmark suites with drift and privacy constraints.
  - Leader boards for synthesis + detection.
  - Open-source code, seeds, and checkpoints.
- **Metrics**: Fidelity, downstream utility, privacy leakage, reproducibility.

Ethical & Societal Guardrails

- **Watermark synthetic data** to ensure provenance.
- **Audit privacy leakage** before deployment.
- **Mitigate dual-use risks** with adversarial testing and policy frameworks.

Hybrid generative models

Hybrid Generative Models are like the ensemble cast of a blockbuster: each model brings its own strengths, and together they can tackle challenges that solo models struggle with.

🔍 Why Hybrid Generative Models Matter

Each generative technique has its own superpowers — and blind spots:

Model Type	Strengths	Limitations
GANs	Sharp, realistic samples	Mode collapse, unstable training
VAEs	Structured latent space, easy sampling	Blurry outputs, limited expressiveness
Diffusion Models	High fidelity, stable training	Slow sampling, high compute cost
Autoregressive Models	Great for sequential data	Slow generation, limited global context

By combining them, hybrid models can:

- **Balance fidelity and diversity** (e.g., GAN + VAE)
- **Speed up sampling** (e.g., Diffusion + Autoregressive)
- **Improve anomaly detection** by modelling both likelihood and reconstruction error
- **Handle multimodal IoT data** with tailored sub-models for each modality

Example Architectures

- **VAE-GAN**: Uses VAE for latent encoding and GAN for realistic output — great for anomaly detection with interpretable latent space.
- **Diffusion-VAE**: VAE provides a structured latent space, while diffusion refines sample quality.
- **Autoregressive-GAN**: Autoregressive model captures temporal dependencies, GAN sharpens outputs.
- **Multi-branch hybrids**: Different generative heads for different modalities (e.g., sensors vs logs vs images).

Applications in IoT

- **Data Synthesis**: Generate realistic sensor sequences with temporal and spatial coherence.
- **Anomaly Detection**: Use hybrid likelihood + reconstruction-based scoring for robust detection.
- **Representation Learning**: Learn disentangled features for downstream tasks like forecasting or classification.

Evaluation Strategies

- **Fidelity**: Fréchet Distance, MMD, spectral similarity
- **Utility**: Train-on-synthetic/test-on-real performance
- **Anomaly Detection**: AUC-ROC, latency, false alarms
- **Efficiency**: Sampling speed, memory footprint

Hybrid Generative Model for Healthcare IoT

A hybrid generative model for healthcare IoT is a powerful architectural approach that integrates multiple AI paradigms, such as deep generative modeling, federated learning, and privacy-preserving techniques, to enable safe, explicable, and adaptable healthcare solutions across distributed IoT environments. This is a systematic categorization according to your goals and interests.

- **Conceptual Framework**: Variational Autoencoders (VAEs) or Diffusion Models for anomaly detection and synthetic data creation are combined in the Hybrid Generative Model.

Decentralized training across edge IoT devices using federated learning Homomorphic encryption and differential privacy for safe data exchange Clinical decision-making interpretability with Explainable AI (XAI) modules

Key Features

- **Multimodal fusion**: Combines vitals, logs, and device metadata.
- **Temporal + spatial modelling**: CNN + LSTM for rich feature extraction.
- **Generative counterfactuals**: Diffusion model simulates rare health events.
- **Privacy-preserving**: Federated learning + DP ensures data security.
- **explain ability**: Highlights root causes and sensor contributions.

Explainable generative AI

Reasonable The goal of the developing multidisciplinary discipline of generative artificial intelligence (Gen XAI) is to make the internal operations and outputs of generative models, such as diffusion models, big language models, and VAEs, transparent, interpretable, and reliable. It's where accountability and creativity collide.

Generative AI uses patterns discovered in data to produce new text, image, audio, and code. However, these models' choices and results frequently become ambiguous as they become more intricate. Reasonable The goal of generative AI is to provide answers to queries such as: Why did the model produce this result?

What information affected this outcome?

🔍 Key Dimensions of Explainability in GenAI

Dimension	Description
Verifiability	Can the output be traced back to reliable sources or training data?
Interactivity	Can users probe or modify the generation process to understand it better?
Transparency	Are the model's architecture and decision paths interpretable?
Security & Privacy	Are explanations privacy-preserving and robust against misuse?
Cost-awareness	Do explanations balance clarity with computational efficiency?

- **Real-World Application:** Medical Consider a generative model that mimics the course of a disease or synthesizes patient reports. GenXAI would: Emphasize the symptoms or biomarkers that affected the diagnosis that was produced.

Federated and privacy-preserving approaches

By keeping data local, secure, and compliant, federated and privacy-preserving approaches are revolutionizing the way we train and implement machine learning models, particularly in delicate fields like healthcare, finance, and the Internet of Things. This is a well-organized summary of explainable, moral AI for regulated settings that is suited to your interests:

🧠 Core Concepts

Approach	Description
Federated Learning (FL)	Decentralized training across devices/institutions without sharing raw data
Differential Privacy (DP)	Adds noise to model updates to obscure individual data contributions
Secure Multiparty Computation (SMPC)	Enables joint computation without revealing private inputs
Homomorphic Encryption (HE)	Allows computation on encrypted data without decryption
Trusted Execution Environments (TEE)	Hardware-based isolation for secure model training and inference

Example of IoT in Healthcare

Consider several hospitals working together to train a diagnostic model without exchanging patient data:

- FL makes certain that every hospital receives local training.
- DP safeguards patient updates for each individual.
- Aggregation of model weights is secured by SMPC/HE.
- The XAI layer provides clinicians with explanations of predictions.
- Model updates for blockchain logs for auditability

This configuration allows for robust, explainable AI while supporting GDPR, HIPAA, and other compliance frameworks.

Edge AI integration

By combining edge computing and artificial intelligence, edge AI integration enables intelligent decision-making to be delivered straight to sensors, wearables, cameras, and gateways. Real-time, secure, and scalable intelligence across distributed systems is made possible by these devices, which analyze and act on data locally rather than sending it to the cloud for processing.

Key Components of Edge AI

Component	Role in Integration
Edge Devices	IoT sensors, smartphones, cameras, and embedded systems that collect and process data
AI Models	Lightweight or compressed models (e.g., MobileNet, TinyML, quantized transformers)
Edge Hardware	Specialized chips like NVIDIA Jetson, Google Edge TPU, ARM Cortex-M
Model Optimization	Techniques like pruning, quantization, and distillation to fit models on constrained hardware
Local Inference Engine	Frameworks like TensorFlow Lite, ONNX Runtime, or PyTorch Mobile for on-device execution

Benefits of Edge AI Integration

- **Low Latency**: Instant decisions—critical for healthcare alerts, autonomous vehicles, and industrial automation
- **Enhanced Privacy**: Data stays local, reducing exposure and regulatory risk
- **Offline Capability**: Devices operate independently of cloud connectivity

- **Bandwidth Efficiency**: Only essential insights are transmitted
- **Scalability**: Intelligence distributed across thousands of devices

Healthcare IoT Example

Imagine a wearable device monitoring cardiac rhythms:

- Detects arrhythmias locally using a compressed CNN
- Flags anomalies in real time
- Sends only critical alerts to clinicians
- Participates in federated learning across hospitals
- Preserves privacy via differential privacy and secure aggregation

This setup supports GDPR/HIPAA compliance while enabling robust, explainable AI at the edge.

Challenges & Research Frontiers

- **Energy Efficiency**: Balancing inference speed with battery life
- **Model Robustness**: Handling noisy, multimodal sensor data
- **Security**: Protecting models and data from edge-based attacks
- **Interoperability**: Integrating across diverse hardware and protocols

Standardized Benchmarks

Standardized benchmarks are essential for evaluating and comparing AI systems in a **consistent, transparent, and meaningful** way—especially in high-stakes domains like healthcare, IoT, and regulated industries.

What Are Standardized Benchmarks?

Benchmarks are curated datasets and evaluation protocols used to:

- **Measure model performance** across tasks (e.g., diagnosis, anomaly detection)
- **Compare models** under consistent conditions
- **Ensure reproducibility** and fairness in research and deployment
- **Support regulatory compliance** by aligning with safety and ethical standards.

Conclusion

Generative AI is emerging as a transformative force in IoT ecosystems, enabling both **synthetic data generation** and **robust anomaly detection** across distributed, privacy-sensitive environments. By learning complex data distributions, models like VAEs, GANs, and diffusion architectures can simulate realistic sensor patterns, fill gaps in sparse datasets, and detect deviations that signal faults or cyber threats—even zero-day attacks.

In industrial and healthcare IoT, this dual capability addresses two persistent challenges:

- **Data scarcity and imbalance**, which hinder traditional anomaly detectors
- **Privacy and compliance constraints**, which limit centralized data access

When integrated with **federated learning**, **edge AI**, and **privacy-preserving techniques** (e.g., differential privacy, homomorphic encryption), generative models become part of a **modular, explainable, and scalable architecture** for real-time intelligence. This fusion supports proactive fault detection, adaptive system optimization, and ethical deployment in regulated domains.

Ultimately, generative AI doesn't just enhance IoT analytics—it redefines it. It enables systems that are not only intelligent but also **resilient, transparent, and human-aligned**.

References

Aung, Y. L., Christian, I., Dong, Y., Ye, X., Chattopadhyay, S., & Zhou, J. (2025). *Generative AI for Internet of Things security: Challenges and opportunities*. arXiv. https://arxiv.org/abs/2502.08886

Manoj, H., Gadiyar, T., Thyagaraju, G. S., Preethi, B., Shivathaya, K., & Balapradeep, K. N. (2024). *The role of generative AI in the Internet of Things*. In V. Ramesh & S. Kumar (Eds.), Futuristic Trends in IoT (Vol. 3, Book 2, Chapter 29, pp. 365–369). IIP Series. https://doi.org/10.2478/ie-2024-0003

Deng, Z., Torim, A., Ben Yahia, S., & Bahsid, H. (2025). Generative AI in intrusion detection systems for Internet of Things: A systematic literature review. *IEEE Open Journal of the Communications Society, 6*, 4689–4717. https://doi.org/10.1109/OJCOMS.2025.3573194

Cherukuvada, S., Kirthika, R., Bhagyalakshmi, A., Danda, R. R., Vani, N., & Roseline, R. (2025). *Generative AI meets IoT: Transformative applications and emerging paradigms*. SSRN. https://ssrn.com/abstract=5080682

Muthusamy Chinnan, S. P., Drury-Grogan, M., & Chakravarthi, B. R. (2025). Gender inclusive language generation framework: A reasoning approach with RAG and CoT. *Knowledge-Based Systems, 328*, 114092. https://doi.org/10.1016/j.knosys.2025.114092

Shahin, M., Hosseinzadeh, A., & Chen, F. F. (2025). A two-stage hybrid federated learning framework for privacy-preserving IoT anomaly detection and classification. *IoT, 6*(3), 48. https://doi.org/10.3390/iot6030048

Riaz, R., Han, G., Shaukat, K., Khan, N. U., Zhu, H., & Wang, L. (2025). A novel ensemble Wasserstein GAN framework for effective anomaly detection in industrial internet of things environments. *Scientific Reports, 15*, Article 26786. https://doi.org/10.1038/s41598-025-07533-1

Rhachi, H., Balboul, Y., & Bouayad, A. (2025). Enhanced anomaly detection in IoT networks using deep autoencoders with feature selection techniques. *Sensors, 25*(10), 3150. https://doi.org/10.3390/s25103150

López Delgado, J. L., & López Ramos, J. A. (2024). A comprehensive survey on generative AI solutions in IoT security. *Electronics, 13*(24), Article 4965. https://doi.org/10.3390/electronics13244965

Liu, Y., Liu, J., Li, C., Xi, R., Li, W., Cao, L., Wang, J., Yang, L. T., Yuan, J., & Zhou, W. (2025). Anomaly detection and generation with diffusion models: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Advance online publication. https://arxiv.org/abs/2506.09368

Lim, W., Yong, K. S. C., Lau, B. T., & Tan, C. C. L. (2024). Future of generative adversarial networks (GAN) for anomaly detection in network security: A review. *Computers & Security, 139*, 103733. https://doi.org/10.1016/j.cose.2024.103733

Afrin, S., Rafa, S. J., Kabir, M., Farah, T., Bin Alam, M. S., Lameesa, A., Ahmed, S. F., & Gandomi, A. H. (2025). Industrial Internet of Things: Implementations, challenges, and potential solutions across various industries. Computers in Industry, 170, 104317. https://doi.org/10.1016/j.compind.2025.104317. *https://doi.org/10.1016/j.compind.2025.104317*

Pandiyan, P., Saravanan, S., Kannadasan, R., Krishnaveni, S., Alsharif, M. H., & Kim, M.-K. (2024). A comprehensive review of advancements in green IoT for smart grids: Paving the path to sustainability. *Energy Reports, 11*, 5504–5531. https://doi.org/10.1016/j.egyr.2024.05.021

Allioui, H., & Mourdi, Y. (2023). Exploring the full potentials of IoT for better financial growth and stability: A comprehensive survey. *Sensors, 23*(19), 8015. https://doi.org/10.3390/s23198015

Sadanand, V. K. (2019). IoT applications in finance and banking. *International Journal of Research and Analytical Reviews (IJRAR), 6*(2), 951–954. https://doi.org/10.6084/m9.doi.one.IJRAR19K3720

Kalsoom, T., Ramzan, N., Ahmed, S., Anjum, N., Safdar, G. A., & Ur Rehman, M. (2025). Socio-organisational challenges and impacts of IoT: A review in healthcare and banking. *Journal of Sensor and Actuator Networks, 14*(3), 46. https://doi.org/10.3390/jsan14030046.

Gugueoth, V., Safavat, S., Shetty, S., & Rawat, D. (2023). A review of IoT security and privacy using decentralized blockchain techniques. *Computer Science Review, 50*, 100585. https://doi.org/10.1016/j.cosrev.2023.100585

Matthew, P., Mchale, S., Deng, X., Nakhla, G., Trovati, M., Nnamoko, N., Pereira, E., Zhang, H., & Raza, M. (2025). A review of the state of the art for the Internet of Medical Things. *Sci, 7*(2), 36. https://doi.org/10.3390/sci7020036

Kabeyi, M. J. B., & Olanrewaju, O. A. (2023). Smart grid technologies and application in the sustainable energy transition: A review. *International Journal of Sustainable Energy, 42*(1), 685–758. https://doi.org/10.1080/14786451.2023.2222298

Zaman, M., Puryear, N., Abdelwahed, S., & Zohrabi, N. (2024). A review of IoT-based smart city development and management. *Smart Cities,* 7(3), 1462–1501. https://doi.org/10.3390/smartcities7030061

Geiger, A., Cuesta-Infante, A., Alnegheimish, S., Liu, D., & Veeramachaneni, K. (2020). Time series anomaly detection using generative adversarial networks (TadGAN). *arXiv preprint arXiv:2009.07769*. https://arxiv.org/abs/2009.07769

Liu, J., Ma, Z., Wang, Z., Zou, C., Ren, J., Wang, Z., Song, L., Hu, B., Liu, Y., & Leung, V. C. M. (2025). A survey on diffusion models for anomaly detection. *arXiv*. https://arxiv.org/abs/2501.11430

Chen, Z., Chen, D., Zhang, X., Yuan, Z., & Cheng, X. (2022). Learning graph structures with transformer for multivariate time series anomaly detection in IoT. *arXiv*. https://arxiv.org/abs/2104.03466

Patel, K. K., & Patel, S. M. (2016). *Internet of Things-IOT: Definition, characteristics, architecture, enabling technologies, application & future challenges*. International Journal of Engineering Science and Computing, 6(5), 6122–6131. https://doi.org/10.4010/2016.1482

Li, Y., Chen, W., Ding, Y., Qie, Y., & Zhang, C. (2022). A vision of intelligent IoT trends, characteristics and functional architecture. In *2022 International Wireless Communications and Mobile Computing Conference (IWCMC)* (pp. 184–189). IEEE. https://doi.org/10.1109/IWCMC55113.2022.9824304

Shankesh, R. M. (2022). IoT data management in Oracle databases: Challenges and solutions. *International Journal for Multidisciplinary Research*, 4(6), 1–18. https://doi.org/10.36948/ijfmr.2022.v04i06.38925

Shaukat, K., Alam, T. M., Hameed, I. A., Khan, W. A., Abbas, N., & Luo, S. (2021). A review on security challenges in Internet of Things (IoT). In *Proceedings of the 26th International Conference on Automation & Computing* (pp. 1–8). University of Portsmouth. https://doi.org/10.1109/IAC.2021.00000

Zikria, Y. B., Ali, R., Afzal, M. K., & Kim, S. W. (2021). Next-generation Internet of Things (IoT): Opportunities, challenges, and solutions. *Sensors*, 21(4), 1174. https://doi.org/10.3390/s21041174

Ruan, Z. (2023). Blockchain technology for security issues and challenges in IoT. In *2023 International Conference on Computer Simulation and Modeling, Information Security (CSMIS)* (pp. 572–580). IEEE.

https://doi.org/10.1109/CSMIS60634.2023.00108

Brühl, V. (2024). Generative artificial intelligence – Foundations, use cases and economic potential. *Intereconomics*, 59(1), 5–9. https://doi.org/10.2478/ie-2024-0003

Narapareddy, V. S. R. (2025). *Generative AI and foundation models*. Universal Library of Innovative Research and Studies, 2(2), 7–21.

https://doi.org/10.70315/uloap.ulirs.2025.0202002

Dash, A., Yer, J., & Wang, G. (2024). A review of generative adversarial networks (GANs) and its applications in a wide variety of disciplines: From medical to remote sensing. *IEEE Access*, 12, 18330–18357.
https://doi.org/10.1109/ACCESS.2023.3346273.

☯☯☯