IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 12 April 2024, accepted 3 May 2024, date of publication 6 May 2024, date of current version 24 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3397775

TOPICAL REVIEW

Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers

STAPHORD BENGESI[1], (Graduate Student Member, IEEE), HODA EL-SAYED[1], MD KAMRUZZAMAN SARKER[1], YAO HOUKPATI[1], JOHN IRUNGU[2], AND TIMOTHY OLADUNNI[3]

[1]Department of Computer Science, Bowie State University, Bowie, MD 20715, USA
[2]Department of Computer Science, University of the District of Columbia, Washington, DC 20008, USA
[3]Department of Computer Science, Morgan State University, Baltimore, MD 21251, USA

Corresponding author: Staphord Bengesi (sbengesi@bowiestate.edu)

ABSTRACT The launch of ChatGPT in 2022 garnered global attention, marking a significant milestone in the Generative Artificial Intelligence (GAI) field. While GAI has been in effect for the past decade, the introduction of ChatGPT sparked a new wave of research and innovation in the Artificial Intelligence (AI) domain. This surge has led to the development and release of numerous cutting-edge tools, such as Bard, Stable Diffusion, DALL-E, Make-A-Video, Runway ML, and Jukebox, among others. These tools exhibit remarkable capabilities, encompassing tasks ranging from text generation and music composition, image creation, video production, code generation, and even scientific work. They are built upon various state-of-the-art models, including Stable Diffusion, transformer models like GPT-3 (recent GPT-4), variational autoencoders, and generative adversarial networks. This advancement in GAI presents a wealth of exciting opportunities across various sectors, such as business, healthcare, education, entertainment, and media. However, concurrently, it poses unprecedented challenges such as impersonation, job displacement, privacy breaches, security vulnerabilities, and misinformation. To addressing these challenges requires a new direction for research to develop solutions and refine existing products. In our endeavor to contribute profound insights to society and advance research on GAI, we present a comprehensive journal which explores the theoretical and mathematical foundations of GAI state-of-the-art models, exploring the diverse spectrum of tasks they can perform, examining the challenges they entail, and discussing the promising prospects for the future of GAI.

INDEX TERMS Generative AI, GPT, bard, ChatGPT, diffusion model, transformer, GAN, autoencoder, artificial intelligence.

I. INTRODUCTION

The release of ChatGPT on November 30, 2022 [1], [2], triggered an exponential surge in the groundbreaking and widespread popularity of GAI to the general public. This remarkable achievement could be traced to the 1956 summer project at Dartmouth College spearheaded by McCarthy; marking the inception of artificial intelligence (AI) [3]. The endeavor aimed to develop machines with the ability to perform tasks typically demanding human intelligence [4], [5], [6], [7], [8]. These tasks include computer vision, natural language processing (NLP) [9], [10], robotics, and many others. Since then, significant advancements have been achieved in imbuing machines with the capability of talking, walking, thinking, and acting like humans. Notably, a series of algorithms, including the Regression model, perceptron algorithm [11], Decision tree [12], K-Nearest Neighbor [13], Naive Bayes Classifier, Back Propagation, support vector machine (SVM) [14], and Random Forest [15] have emerged. These algorithms in contemporary times

The associate editor coordinating the review of this manuscript and approving it for publication was Nikhil Padhi.

are commonly referred to as classical/traditional machine learning algorithms and most of them were developed before the year 2000. Furthermore, there is an advancement in deep learning algorithms, including the development of Convolutional Neural Networks (CNNs) in the 1980s [16], Recurrent Neural Networks (RNNs) in 1985 [17], Long Short-Term Memory (LSTM) in 1997 [18], and Bidirectional Long Short-Term Memory (BiLSTM) [19] in the same year. However, until recent times, widespread attention has been limited primarily because of computing resources and dataset availability limitations [20].

To tackle the constraints imposed by limited datasets, researchers from Stanford University, Princeton University, and Columbia University jointly launched the ImageNet Large Scale Visual Recognition Challenge in 2010 [21]. This competition played a pivotal role in driving advancements in neural network architectures, with a particular focus on Convolutional Neural Networks (CNNs). Since then, CNN has been established as an algorithm for image classification and computer vision [22]. The breakthrough achievement of AlexNet in 2012 [23] marked a significant milestone in the practical application of deep learning in computer vision tasks. The success of the ImageNet Competition ignited a surge in interest and investment in deep learning research. This newfound enthusiasm resulted in the continuous evolution of improved architectural innovations, including models such as ResNet [24], DenseNet [25], MobileNet [26], and EfficientNet [27]. These models set the gold standard for various cutting-edge technologies, such as transfer learning, continual learning, attention mechanisms [28], self-supervised learning, and generative AI.

Before 2014, all existing deep learning models were primarily descriptive, focusing on summarizing or representing existing data patterns and relationships. These models aimed to explain the data patterns and make predictions based on the information present. However, Goodfellow et al. [29] in 2014 introduced the Generative Adversarial Network (GAN) ushering in a new era of Generative Artificial Intelligence (GAI) realization. Unlike their descriptive counterparts, generative models, such as GANs, are designed to learn the underlying probability distribution of the data [30]. Their primary goal is to generate new data samples that closely resemble the patterns observed in the training data [31], [32].

The breakthrough of GAN marked a significant departure from traditional deep learning methods, opening exciting possibilities for Generative artificial intelligence. GAI has since garnered widespread attention due to its transformative impact across various domains of life. It offers elegant solutions to complex problems [33] enabling the creation of synthetic data, artistic content, and realistic simulations. This paradigm shift in AI technology has profoundly influenced the new perception, implementation, and utilization of artificial intelligence, sparking innovation and new application opportunities across industries.

The rise of GAI has sparked countless inquiries, prompting the necessity for comprehensive GAI exploration. Despite numerous recent studies to address the surge of GAI [31], [31], [34], [34], [35], [35], [36], [36], [37], [38], [39], [40], [41], [44], [45], [46], discussing challenges, tasks, and models, there still needs to be more thorough exploration into the theoretical and mathematical foundations of the recent GAI models and their related aspects such tools which are evolving exponentially as schematically presented on Figure 1. Motivated by this observation, our goal is to fill this gap by conducting an in-depth review of the state-of-the-art in GAI. Our contribution involves scrutinizing the most commonly used models, delving into their technical and mathematical backgrounds, listing their latest associated end products (tools), describing task categorization, applications, areas of impact, challenges, and prospects. Through this study, we aim to provide the broader public and new researchers with profound insights into GAI and facilitate its advancement as we enter a new era of advanced AI.

The rest of the work is structured as follows: Section II introduces contemporary generative models. Section III elaborates on the various tasks within GAI. Section IV examines the diverse applications of GAI. Section V delves into the outlook for GAI. Lastly, Section VI offers a conclusion.

II. GENERATIVE MODELS

There has been a shift in the focus of researchers from discriminative learning to generative learning in the contemporary era. Multiple generative models have emerged with the capability of generating new data points like the training data inputs based on learning their distribution. This section will discuss current state-of-the-art theoretical and mathematical foundations of generative models.

A. AUTOENCODER

Autoencoder (AE) is an unsupervised machine learning neural network model that encodes the input data using an encoder into a lower-dimensional representation (encoding) and then uses a decoder to decode it back to its original form (decoding) while reducing the reconstruction error [47]. This model was primarily designed for Dimensionality Reduction, Feature Extraction, Image Denoising, Image Compression, Image Search, Anomaly Detection and Missing Value Imputation [47].

Both encoder and decoder of the model are neural networks written as a function of input and a generic function of code layer respectively [48]. Based on Figure 2, autoencoder is made up of four components namely:

- **Encoder:** This component reduces and compresses the input data into lower dimensions. As a result of its output, it creates a new layer called code.
- **Code/Bottleneck:** a layer that contains a compressed and the lowest possible dimensions of input data representation. Consider equation 1 below.

$$h_i = f\,(X_i) \tag{1}$$

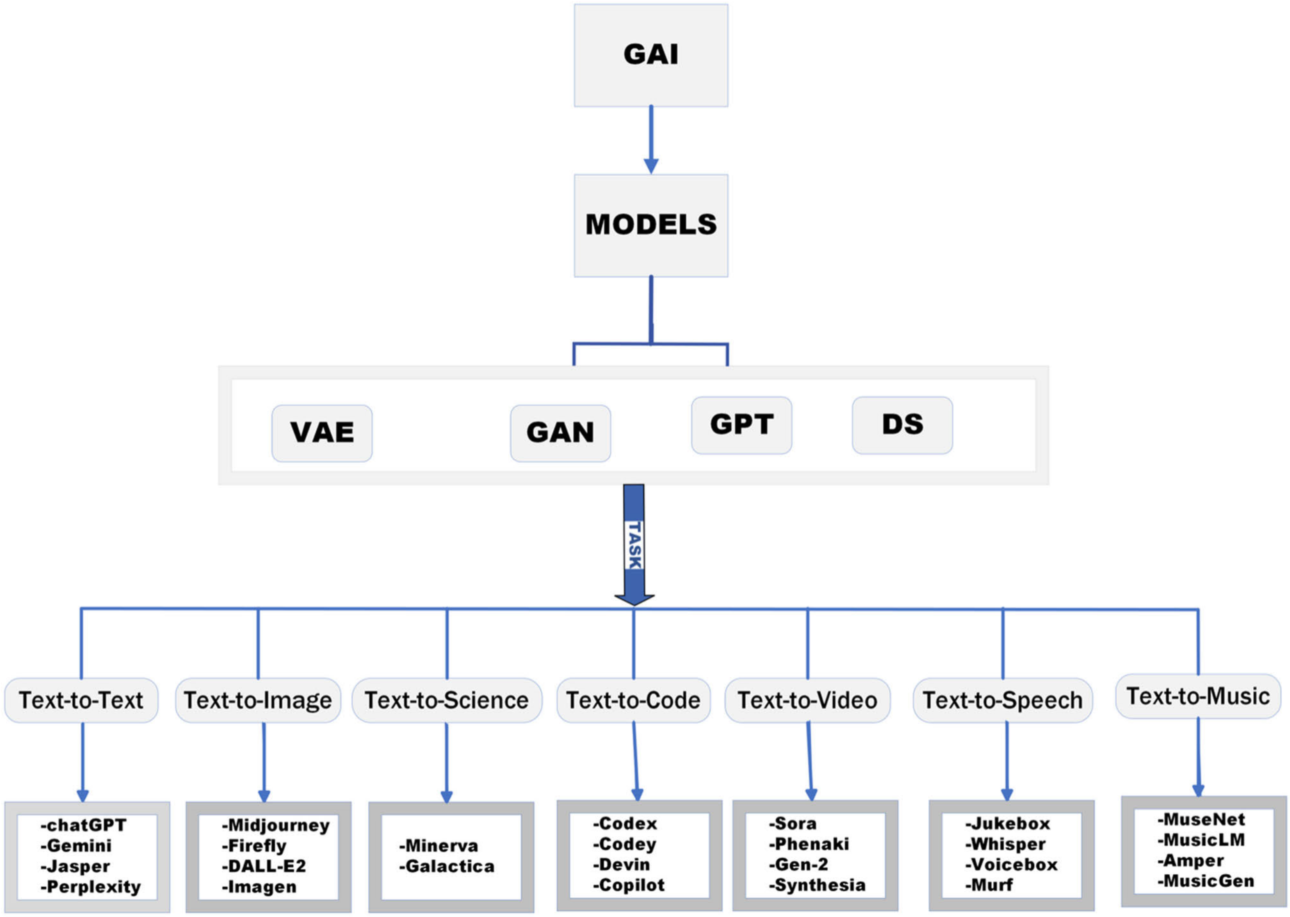


FIGURE 1. The schematic overview of GAI, including a comprehensive list of the most used models, their tasks, and tools use cases.

whereby $\boldsymbol{h_i}$ is code layer after function $\boldsymbol{f}$ *with* user defined parameters is applied to the input $\boldsymbol{X_i}$

- **Decoder**: Reconstructs the code layer from lower dimension representation to input.
- **Reconstruction Loss**: Defines the final output of the decoder, measuring how closely the output resembles the original input.

$$\tilde{X}_i = g\left(h_i\right) \tag{2}$$

where $\tilde{X}_i$ is the output of encoder after second generic function to the code layer.

The training of the autoencoder involves minimizing the dissimilarity between the input and the output [48], as shown in Equation 3.

$$Argmin_{f,g} < \Delta\left(\boldsymbol{X_i}, \tilde{X}_i\right) \tag{3}$$

The encoder and the decoder are composed of fully connected feedforward neural networks where the input, code, and output layers consist each of a single neural network layer defined by the user. Like other standard neural networks, autoencoders apply activation functions such as sigmoid and Relu. Various variants of autoencoder exist, such as contractive, Denoising, and sparse autoencoder [49]. Generally, the plain autoencoders prior mentioned are not

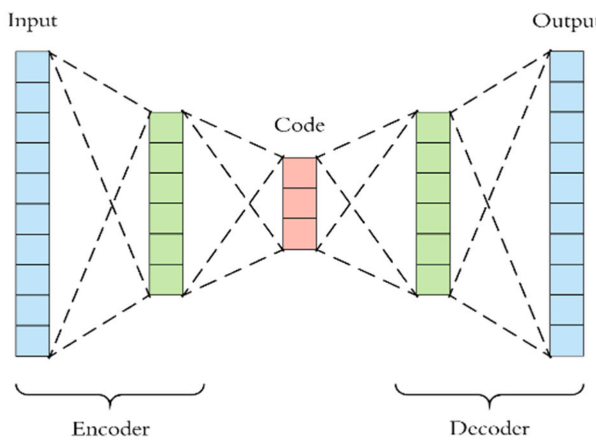


FIGURE 2. Autoencoder architecture[1] describing the main components of AE such as encoder, decoder, and code.

[1]Source: https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798.

TABLE 1. VAE state-of-the-art.

category	Subcategory Domains	Dataset	References
Image Processing	Image Classification	MRI datasets, SAR images, ImageNet dataset, NWPU-RESISC45	[52] [53] [54] [55]
	Image Compression	Kodak dataset	[56]
	Image Resolution	DIV2K and Flickr2K image dataset	[57]
Audio Processing	Noisy voice recorded datasets	NIL	[58] [59]
Video Processing	Prediction	MineRL and MMNIST	[60]
Video	Infrastructure Monitoring	UCSD Ped2, Fan deterioration simulated dataset	[61] [62] [63]
Audio	Infrastructure Monitoring	MIMII	[64]
Sensor	Nonlinear analysis	Simulated Dataset, Butane content	[65] [66]
	Modeling	DCS	[67] [68] [69]
Activity monitoring	Finance	284,807 credit card transactions	[70] [71]

generative since they do not generate new data but replicate the input. However, the variational autoencoder is the variant that is generative [47].

1) VARIATIONAL AUTOENCODER

Variational autoencoder (VAE) evolved as a result of the introduction of variational inference (A statistical technique for approximating complex distributions) to Autoencoder (AE) by Kingma et al. [50]. It's a generative model that utilizes Variational Bayes Inference to describe data generation using a probabilistic distribution [51].

Unlike traditional AEs, VAEs have an extra sampling layer in addition to an encoder and decoder layer as depicted in Figure 3. Training the VAEs model involves encoding the input as a distribution over the latent space and generating the latent vector from the distribution sampling. Afterward, the latent vector is decoded, the reconstruction error is computed, and the reconstruction error is backpropagated through the network. During the training process, regularization is introduced explicitly to prevent overfitting.

Probabilistically, VAE is composed of a latent representation z, drawn from the prior distribution $\boldsymbol{p(z)}$ and the data $\mathbf{x}$ drawn from the conditional likelihood distribution $\boldsymbol{p(x|z)}$

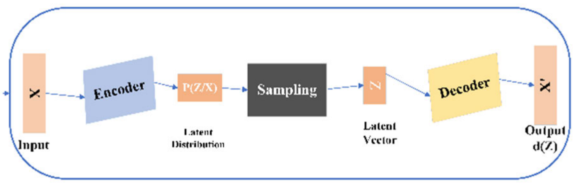


FIGURE 3. The VAE architecture showing its modification from AE incorporating additional layer such as sampling layer, latent distribution, and vector.

which is referred to as probabilistic decoder and can be expressed as:

$$p(x, z) = p(x \mid z)p(z) \tag{4}$$

The inference of the model is examined by computing the posterior of the latent vector using the Bayes theorem shown in equation 5.

$$p(z|x) = \frac{p(x|z)p(z)}{p(x)} \tag{5}$$

With any distribution variant such as Gaussian, variational inference can approximate the posterior, and its reliability in approximation can be assessed through Kullback-Leibler divergence which measures the information lost during approximation. This model has significantly influenced GAI, as demonstrated in Table 1, which highlights a few outstanding state-of-the-art examples using VAE across various domains.

B. TRANSFORMER

The ground-breaking work of Vaswani et al. . *"Attention Is All You Need"* by the Google Brain team introduced a transformer model which can analyze large-scale dataset [28]. Transform was initially developed for NLP but was subsequently adapted to other areas of machine learning, such as computer vision [72], [73], [74]. This model aimed to solve RNNs, and CNNs shortcomings such as long-range dependencies, gradient vanishing, gradient explosion, the need for larger training steps to reach a local/global minima, and the fact that parallel computation was not allowed [28]. Thus, the proposed solution presented a novel way of handling neural network tasks like translation, content generation, and sentiment analysis [75]

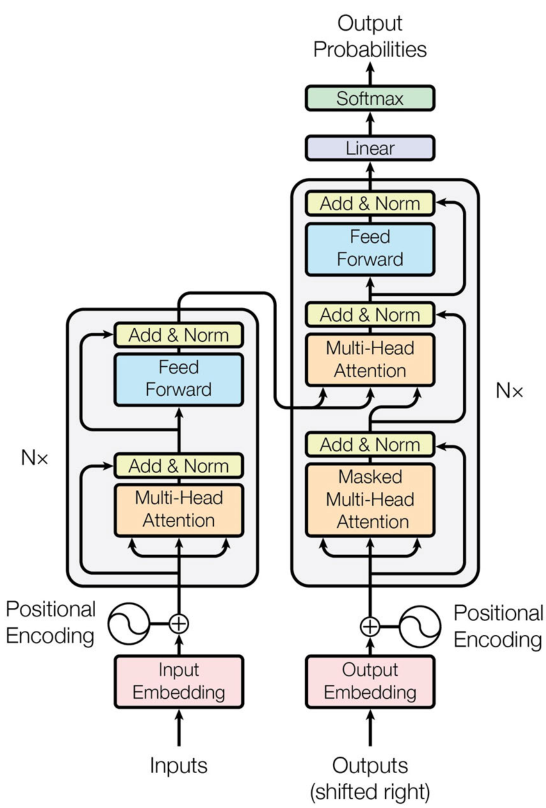


FIGURE 4. The transformer architecture [28] including Self-attention, multi-head attention and word embedding.

Transformer Architecture: Vaswani et al. [28], introduced three main concepts in their study as depicted in Figure 4, including self-attention, which allows a model to evaluate input sequences according to their importance, thus reducing long-range dependencies, multi-head attention which allows the model to learn multiple means of the input sequence, and word embedding, which transforms inputs into vectors.

Encoder and Decoder: It is worth mentioning that the transformer architecture (Figure 3) inherits the encoder-decoder structure [76] that utilizes stacked self-attention and point-wise layers, fully connected layers for both the encoder and decoder [77]. The encoder consists of a stack of N = 6 identical layers, each with two sublayers, including a multi-head self-attention mechanism and a fully connected feedforward network. A decoder is like an encoder, but with an additional sublayer which masks the multi-head attention. Encoders and decoders both apply residual connections to the sublayers, followed by normalization of the layers.

Self-Attention: Attention describes the mechanism for a better understanding of the word's context by paying attention to the vital part of the sentence or any input. It involves mapping a vector of query and a set of key-value pairs to an output vector. According to [28], self-attention refers to Scaled Dot-Product Attention consisting of queries and key dimensions d_k, and dimension d_v values computed according to the following formula:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \tag{6}$$

Figure 5 depicts the structure attention whereby the SoftMax activation function is used to compute the weights on values.

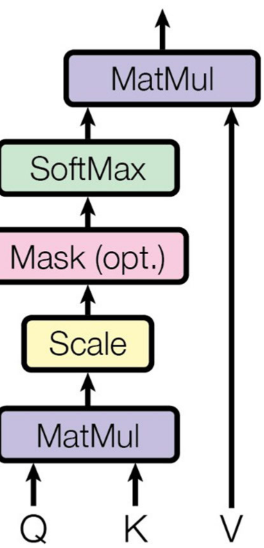


FIGURE 5. Detailed description self-attention architecture [28].

Multi-Head Attention: A multi-head attention mechanism proposes that self-attention can be run multiple times in parallel mode combining knowledge of the same attention pooling via different representation subspaces of queries, keys, and values. Afterward, the independent attention outputs are concatenated and linearly transformed into the expected dimension, as portrayed by equation 5 and Figure 6.

$$MultiHead\ (Q, K, V) = concat\ (head_1, \ldots head_h)\ W^o \tag{7}$$

where $head_i = \text{Attention}\ (QW_i^Q, KW_i^k, VW_i^v)$

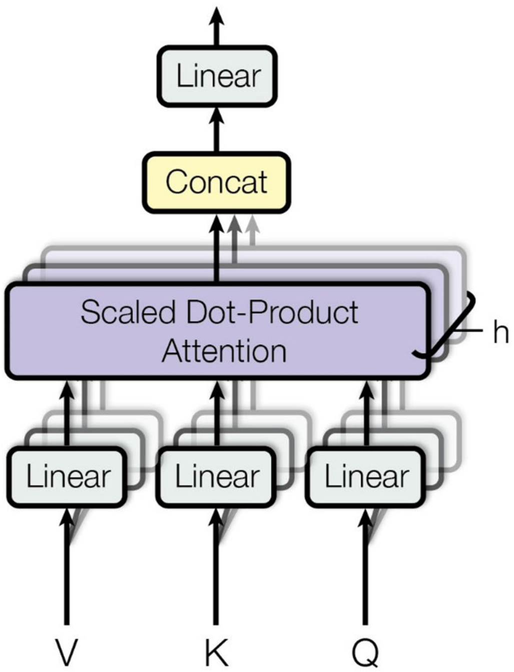


FIGURE 6. The scaled dot product of self-attention architecture [28].

Since the Transformer's invention, several variants have been developed to solve different machine-learning tasks in computer vision and NLP. It's imperative to note that the state-of-the-art models are built on the foundation transformer architecture [72]. In the following subsection, we will discuss the contemporary generative models.

1) GENERATIVE PRE-TRAINED TRANSFORMER (GPT)

A Generative Pretrained Transformer (GPT) describes the transformer-based large language model (LLM) that utilizes deep learning techniques to generate a human-like text [78]. The model was introduced by OpenAI in 2018 [79], following Google's 2017 invention of a transformer. It is made of a stack of transformer decoders. They proposed a model consisting of two stages: learning a high-capacity language model from a large corpus of text and fine-tuning it with labeled data during the discriminative task, as depicted in Figure 7.

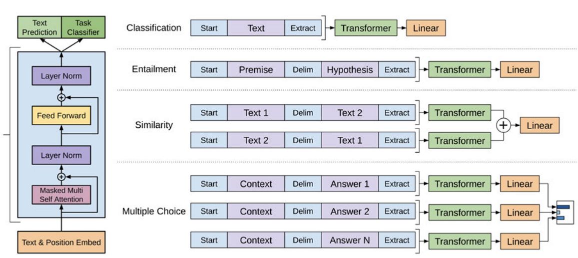


FIGURE 7. The self-attention architecture of GPT [79].

GPT or GPT-1 was trained on the BooksCorpus dataset, which consists of over 7,000 unique unpublished books in many genres, such as Adventure, Fantasy, and Romance, all with long stretches of contiguous text, allowing the generative model to learn on long-range information [61], [62], [65]. The model training specification included the following:

- 12-layer decoder-only transformer.
- Masked self-attention heads (768-dimensional states and 12 attention heads).
- Position-wise feed-forward networks.
- Adam optimization.
- Learning rate: 2.5e-4.
- 3072-dimensional inner states.

The assessment tasks for the model were drawn from four primary categories within NLP: these encompass natural language inference, question answering and common-sense reasoning, semantic similarity, and classification. Following the initial release, OpenAI has produced a series of variant models known as GPT-n series as summarized in table 2, where every successor model is more substantial and efficient than the predecessor. GPT-4 is the most recent variant release in March 2023.

2) GPT-2

After the great success of GPT-1, OpenAI released a second version (GPT-2) in 2019 with 1.5 billion learnable parameters, ten times more in pre-training corpus and parameters than its predecessor trained on WebText, a collection of millions of webpages [80]. As a result, this model can handle complex problems and generate coherent and contextually relevant texts across a wide range of topics and styles.

3) GPT-3

This version was released in 2020 and had 2048-token contexts, 175 billion learnable parameters, which is more than 100 times its predecessor, and required 800GB of storage [81]. CommonCrawl was used to train the model, which was tested on all domains of NLP, and it had promising few-short and zero-shot performance. This version was further improved to GPT 3.5, which was used to develop ChatGPT. Considerable research work has been conducted, incorporating GPT-1 to GPT-3.5 across various task such as Speech Recognition [82], [83], [84], Text Generation [85], [86], [87], [88], [89], [90], [91], [92], Cryptography [93], [94], [95], [96], Computer Vision [97], [98], and Question Answering [99], [100], [101], [102], [103].

4) GPT-4

In March 2023, the most recent GPT model was released by OpenAI [104]. It's a multimodal transformer model, A large-scale language model that accepts image and text inputs and produces text outputs. GPT-4 exhibits high performance comparable to that of humans when tested in several number of professional and academic benchmarks, including passing a bar and medical exam [105], [106]. The model was trained using publicly available internet data and data licensed from third parties and then fine-tuned using Reinforcement Learn-

TABLE 2. GPT's summary.

	GPT-1	GPT-2	GPT-3/ GPT 3,5	GPT-4
Training Parameters	117 million	1.5 billion	175 billion	Unknown
Dataset	BooksCorpus	WebText	CommonCrawl	Public and Private available dataset
Release Date	June 2018	February 2019	GPT-3 was released on June 2020, GPT.3.5 March 2022	March 2023
Maximum Token Length	1024	1024	4096	8192-32,768
NLP tasks	Yes	Yes	Yes	Yes
Image Generation	No	No	No	Yes
Academic and Professional Performance Benchmark	No	No	No	Human Level performance on Bar, Medical, and SAT exam

ing from Human Feedback (RLHF). It was compared with state-of-the-art models using Measuring Massive Multitask Language Understanding (MMLU) [107] that covers 57 tasks in elementary mathematics, US history, computer science, law, and more and outperformed them all.

C. GENERATIVE ADVERSARIAL NETWORK (GAN)

1) GAN OVERVIEW

A generative adversarial network (GAN) is an unsupervised generative model that consists of two neural networks: a generator and a discriminator. A generator attempts to fabricate new data (fake) that is indistinguishable from real data, while a discriminator tries to distinguish between real and fabricated data [108]. Figure 8 illustrate the schematic architecture of GAN (Also known as a vanilla GAN). The generator network takes noise as input and generates fake data. The discriminator network takes both real and fake data as input and classifies them as real or fake using a sigmoid activation function and binary cross-entropy loss [109]. Since the generator does not have direct access to authentic images, it only learns through interactions with the discriminator; the discriminator has access to synthetic and authentic images. Upon completion of classification, backpropagation takes place to optimize the training process [108]. This process repeats itself until the difference between real and fake data samples is negligible.

According to Goodfellow et al. [29], the generator (G) and discriminator (D) are trained together in a minimax game (zero-sum game). In this game as demonstrated by equation 8, G is trying to maximize the probability that D misclassifies its output as real data, while D is trying to minimize the probability that it misclassifies G's output.

$$min_G max_D V(D, G) = E_{x\sim p_{data}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1 - D(G(z)))] \tag{8}$$

where E is the Expected Value, $p_{data(x)}$ is Real data distribution and $p_z(z)$ implies Noise data distribution.

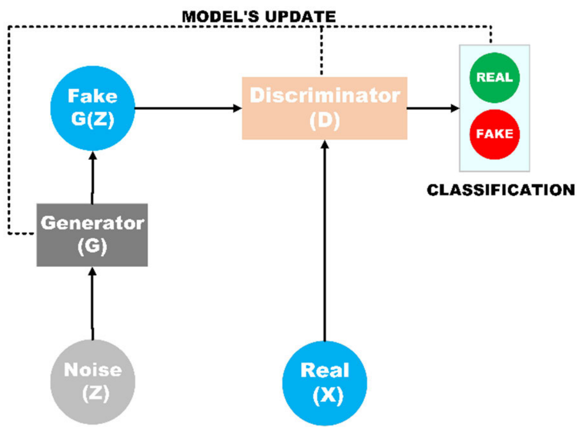


FIGURE 8. The schematic GAN architecture consisting of Discriminator and Generator model.

2) GAN CHALLENGES

Despite their robustness, traditional GANs suffer from limitations such as:

Mode collapse: In this phenomenon, the generator can only produce a single type of output or a limited number of outputs [110]. This is because the generator becomes stuck in a particular mode or pattern, failing to generate diverse outputs that cover the entire data range [111]. There are two main causes of mode collapse in GANs. The first is catastrophic forgetting [112], which occurs when learning in a current task destroys knowledge learned in a previous task. The second cause is discriminator overfitting, which results in the generator loss vanishing [113].

Non-convergence and Instability: The loss function in equation 8 can cause the generator to suffer from gradient vanishing [114]. This can happen when the discriminator learns too quickly and can easily distinguish between real and fake samples. However, the generator may have a lower learning rate and be unable to keep up. This can lead to the training process stalling, as the generator cannot learn from the feedback provided by the discriminator. GANs are also known to be sensitive to the choice of hyperparameters, such

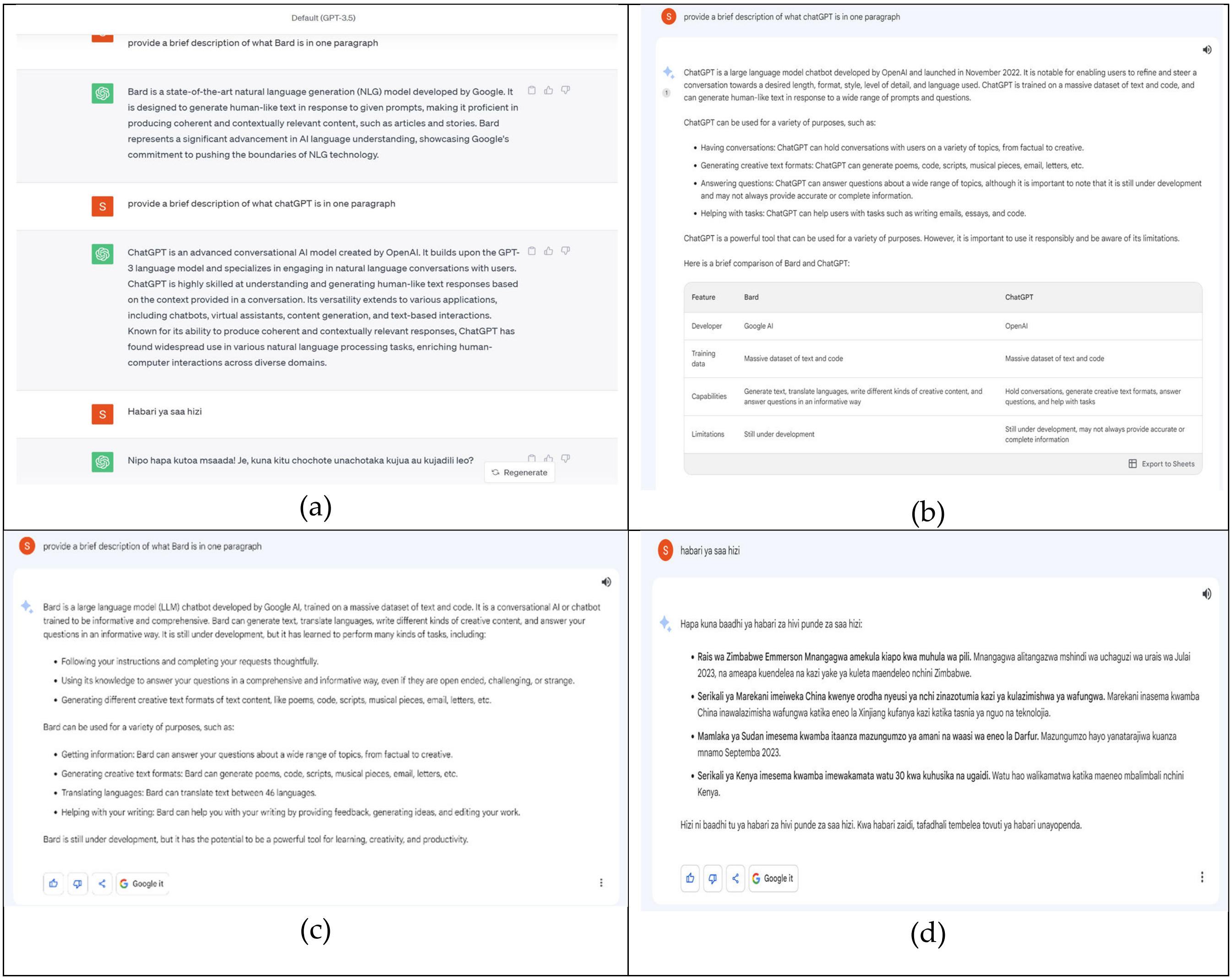


FIGURE 9. **(a) Detailed ChatGPT chatbot and (b)-(d) bard (Now Gemini) outputs for prompts "Briefly describe Bard in one paragraph", "ChatGPT's brief description in one paragraph" and a Swahili question, "Habari za saa hizi".**

as the learning rate and the batch size. This means that it can be challenging to train GANs consistently, as even small changes to the hyperparameters can significantly impact the results [115].

Gradient vanishing can be addressed using a different loss function, such as the Wasserstein loss. The Wasserstein loss is less sensitive to the discriminator's learning rate, and it can prevent the generator's gradients from disappearing. Another solution would be to use a generator with a smaller learning rate. This will prevent generator weights from becoming too large, which can also contribute to gradient vanishing. In addition, a good initialization technique must be used for the generator. In this manner, the generator will start well, and the training process will likely be successful.

3) GAN VARIANTS

In response to the aforementioned GAN challenges, various variants have been developed to address the weaknesses and optimize the model. Here are some of the most famous variants of GAN since its emergence in 2014:

Conditional Generative Adversarial Network (cGAN) cGAN was introduced by Mirza and Osindero [116] in 2014, this variant enhances the classical GAN by incorporating extra auxiliary information into the Generator and Discriminator networks, such as class labels or style attributes. This integration is achieved by introducing an additional layer that includes the conditional information input to the generator, instructing it on what to produce [117]. For instance, in an image generation scenario, this condition might consist of a class label that precisely defines the type of image to be generated.

The Deep Convolutional GAN (DCGAN) framework employs a deep learning model for discriminator and generator components, specifically a Convolutional Neural Network (CNN). In the architectural design defined by

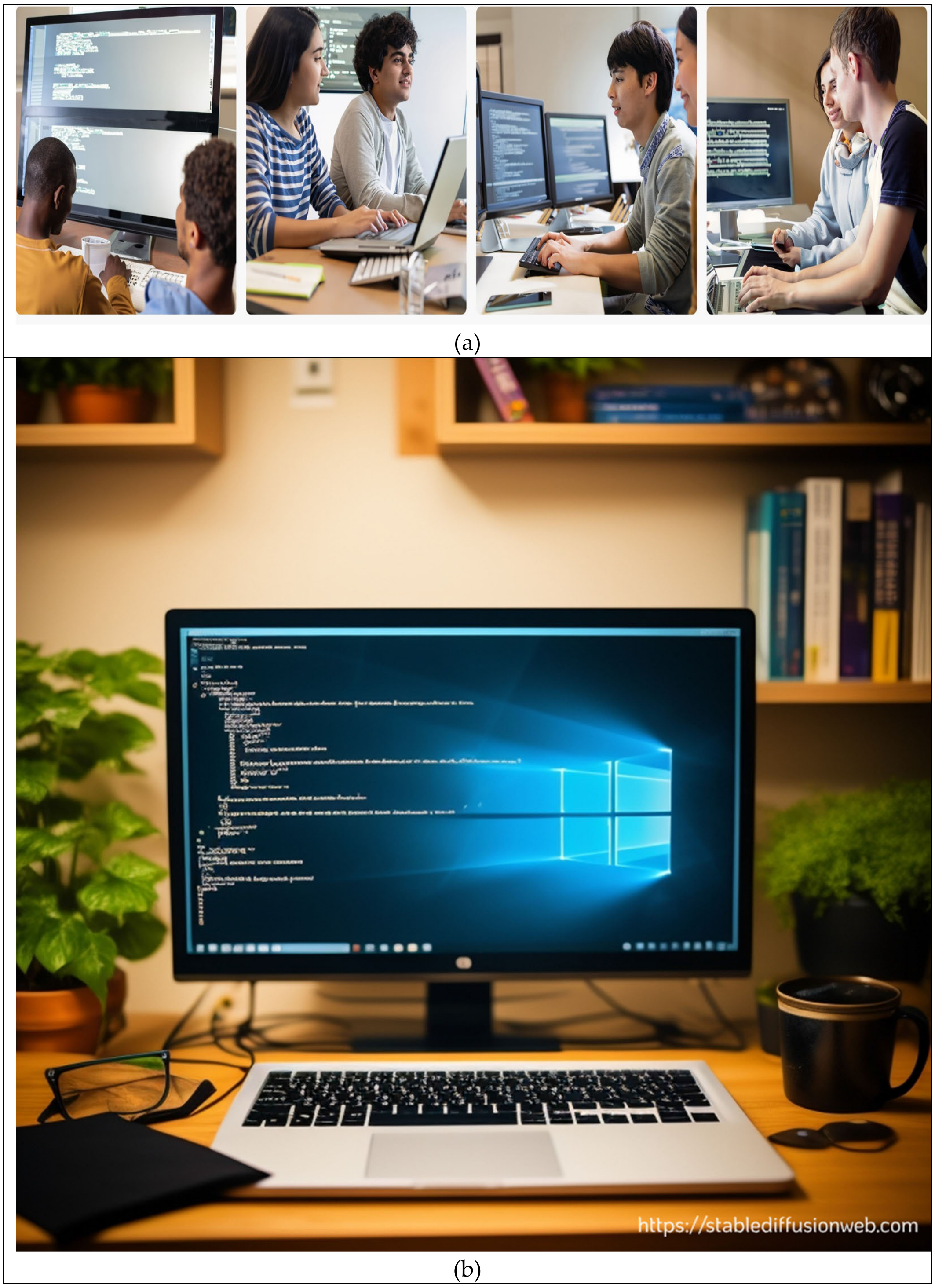


FIGURE 10. **(a) Demonstrate the image generated by the Adobe Firefly tool, while (b) showcase a stable diffusion-generated image both using the "college Student Programming" prompt.**

Radford et al. [118], traditional fully connected layers situated on top of convolutional features have been omitted. Additionally, including Batch Normalization plays a pivotal role in enhancing training stability. This technique normalizes the input to each neural unit, ensuring a mean of zero and unit variance, thus facilitating more consistent and efficient learning. Moreover, DCGAN substitutes conventional pooling layers with strided convolutions in the discriminator and fractional-strided convolutions in the generator network. The Rectified Linear Unit (ReLU) serves as the activation function for the generator, while the Leaky ReLU is employed in the discriminator. These activation functions play a crucial

a b

FIGURE 11. **(b) Depict the generated new living room design using roomGPT from original living room (a) [154].**

a b

FIGURE 12. **(a)The Bowie State University Natural Science Building (a) [126] used by the runway tool to Generate new building design (b).**

role in enabling the networks to capture intricate patterns and features.

Wasserstein GAN (WGAN) is a GAN variant that employs the Wasserstein distance (also referred to as the Earth Mover's distance) as its loss function, distinguishing itself from traditional GANs that typically use the Jensen-Shannon or Kullback-Leibler divergences. The Wasserstein distance (WD) measures the similarity between the distributions of real and generated samples [119]. It is grounded in the solution to a classical optimization problem known as the transportation problem [120]. In this context, suppose there exists several suppliers, each endowed with a certain quantity of goods, tasked with delivering to several consumers, each having a specified capacity limit. Each supplier-consumer pair incurs a cost for transporting a single unit of goods. The transportation problem aims to identify the most cost-efficient allocation of goods from suppliers to consumers.

$$W\left(P_r, P_g\right) = \inf_{\gamma \in \pi(P_r, P_g)} E_{(x,y)\sim\gamma}[||x-y||] \tag{9}$$

WD is expressed by equation 9, $\boldsymbol{P_r}$ *and* $\boldsymbol{P_g}$ denotes the probability distribution of real ad generated sample respectively. The Lipschitz constraint was utilized to impose weight clipping on the discriminator [121]. This measure enhances training stability, mitigating challenges like mode collapse and saturation loss.

Cycle GAN is an approach that automates training image-to-image translation models without requiring paired examples, leveraging GAN architecture [122]. It utilizes unassociated image collections from distinct source and target domains (e.g. Domain X and Domain Y). The model structure comprises two generators: Generator-X crafts images for Domain X, and Generator-Y generates images for Domain Y. Each generator associated with a its corresponding discriminator for binary classification.

This variant incorporates three loss functions: firstly, the cycle consistency losses ensure that translations between domains maintain a coherent loop, returning to their original point; secondly, the adversarial loss pits the Generator against its corresponding Discriminator, with the Generator striving to generate domain-specific images while the Discriminator distinguishes between translated and real samples; and thirdly, the Identity Loss incentivizes the Generator to

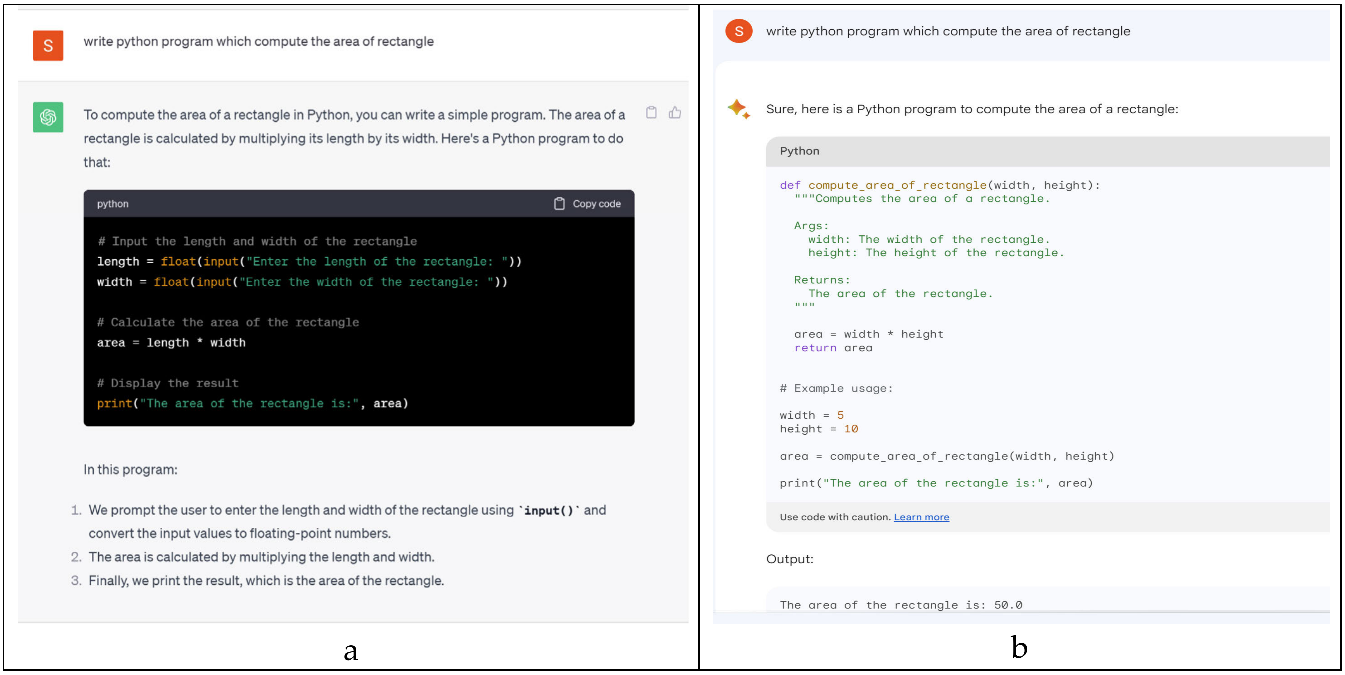


FIGURE 13. Code generation using (a) ChatGPT and (b) Bard tool.

faithfully preserve color composition between input and output, enhancing translation fidelity.

StarGAN: a method that harnesses the power of the GAN architecture for versatile multi-domain image-to-image translation. As outlined by Choi et al [123], this innovative generative adversarial network masterfully learns mappings among numerous domains, employing just a single generator and discriminator, and efficiently trains on images spanning all domains. This model utilizes an Adversarial Loss to make generated images virtually indistinguishable from real ones, a Domain Classification Loss to guarantee precise classification by the discriminator and a Reconstruction Loss that minimizes adversarial and classification losses.

In the preceding subsection, we have delved into several variants of Generative Adversarial Networks (GANs). However, it is worth noting that the landscape of GANs encompasses a myriad of additional variants that have significantly advanced beyond the foundational GAN framework. These notable advancements include the Progressive GAN (PGAN) of 2017 [124], BigGAN of 2018 [125], StyleGAN [126] and StyleGAN 2 [127] of 2019, along with earlier innovations such as InfoGAN [128], Stacked GAN [129], Bidirectional GAN (BiGAN) [130] from 2016.

D. DIFFUSION MODEL

Diffusion model (DM) is a probabilistic generative model characterized by a two-step process. Firstly, the forward diffusion process introduces Gaussian noise into the training data. Then, the reverse diffusion process, known as denoising, gradually reverses the diffusion step by step to generate new sample data [131]. These models effectively overcome challenges encountered in aligning posterior distributions within VAEs, mitigate the inherent instability in the adversarial objectives of GANs by offering a more stable training objective, and addressing the computational burdens associated with Markov Chain methods [132], [133]. DM basically encompasses three primary formulations: denoising diffusion probabilistic models (DDPMs), stochastic differential equations (SDEs), and score-based generative models (SGMs) [134].

E. DENOISING DIFFUSION PROBABILISTIC MODELS (DDPMS)

DDPMs [135], [136]employ the aforementioned processes of DM, whereby the **forward process** introduces noises $\varepsilon \sim N(0,1)$ into data distribution sample $X_0 \sim p(X_0)$, to generate the noisy data distribution or prior distribution, $p\left(X_t \mid X_{t-1}\right)$ using Markov chain as depicted by equation 10 given that index 0 denotes original data and $p\left(X_0\right)$ represent the probability density of the data.

$$p(X_t|X_{t-1}) = N(X_t; \sqrt{1-\beta_t}X_{t-1}, \beta_t I) \tag{10}$$

where I mean identity matrix and β_t entail variance schedule across various diffusion step t. we can now define $\alpha_t = 1 - \beta_t$ from variance schedule and as long as we know the original data X_0 we use a single step to create a noisy data X_t and generate the sample distribution shown in equation 11.

$$P\left(x_t \mid x_0\right) = N(X_t; \sqrt{\bar{\alpha}_t}X_0, \sqrt{1-\bar{\alpha}_t I}) \tag{11}$$

Based on equation 11, we can derive the definition of X_t as shown on equation 12;

$$X_t = \sqrt{\bar{\alpha}_t}X_0, + \sqrt{1-\bar{\alpha}_t}I. \tag{12}$$

TABLE 3. GAI tools.

	Tool	Developer	Task	Year	Additional Description
1	DeepDream [164]	Google	Text-to-Image	2015	Generate psychedelic images
2	WaveNet [165]	DeepMind	Text-to-Speech	2016	Generate realistic speech from text or other audio inputs
3	Sensei [166]	Adobe	Text-to-Image	2016	Generate automative workflow and personalize cunstomer experience
4	Synthesia [167]	Synthesia	Text-to-Video	2018	Generate video from text input
5	Boomy [168]	Boomy	Text-to-Music	2019	Develop music without prior knowledge.
6	AI Art [169]	Nightcafe	Text-to-Image	2019	Generate image from text input
7	StyleGAN [170]	Nvidia	Text-to-Image	2019	Generate realistic and creative image from text prompt
8	MuseNet [153]	OpenAI	Text-to-Music	2019	Generate Music of various genres
9	Descript [171]	Descript	Text-to-Video	2020	Generate video from text input
10	Genny [172]	Lovo	Text-to-Speech, Text-to-Image	2020	Can generate voice over and art image
11	Murf [173]	Murf.Ai	Text-to-Speech	2020	Generate voice-over for Creative content and Presentation
12	flashGPT	Neuroflash	Text-to-Text	2020	A generative Chatbot which use flash
13	Jukebox [154]	OpenAI	Text-to-Speech	2020	Music Generator
14	Synthesys [174]	Synthesys	Text-to-Speech	2020	Create voiceover from text
15	Photosonic [175]	Writesonic	Text-to-Image	2020	Generate image from text input
16	NovelAI [176]	Anlatan	Text-to-Image	2021	Generate images from text input and storywriting.
17	Overdub [177]	Descript's	Text-to-Speech	2021	Creates realistic voice from text
18	Jasper [178]	Jasper	Text-to-Text	2021	Generate Creative Contents
19	GitHub Copilot [179]	Microsoft/GitHub/OpenAI	Text-to-Code	2021	Code Generator and Suggestion
20	Codex [180]	OpenAI	Text-to-Code	2021	Code Generator
21	Read [181]	Read.ai	Speech-to-Text	2021	Virtual meeting Automated summary, transcripts, playback, and highlights on action items, key questions, and real-time engagement
22	Soundful [182]	soundful	Text-to-Music	2021	Create customized music based on individual needs
23	Soundraw [183]	Soundraw Inc	Text-to-Music	2021	Generate Music
24	StarryAI [184]	StarryAI Inc	Text-to-Image	2021	Generate image from text input
25	TexTalky [185]	Textalky	Text-to-Speech	2021	Creates realistic voice from text
26	Generate [186]	Cohere	Text-to-Text	2022	Content Generation
27	AlphaCode [187]	DeepMind	Text-to-Code	2022	Generate code, creative content, and respond to questions in an informative way
28	Minerva [162]	Google	Text-to-Science	2022	Solve Quantitative reasoning problem
29	LaMDA 2 [188]	Google	Text-to-Speech	2022	Customer Service Chatbots, Q&A, Translation, Research
30	Imagen Video [189]	Google	Text-to-Video	2022	1280x768 HD videos at 24 frames per second from text limited to inanimate objects
31	Galactica [163]	Meta AI	Text-to-Science	2022	tool for scientific writing
32	PEER [190]	Meta AI	Text-to-Text	2022	Writing tool
33	Make-A-Video [148]	Meta AI	Text-to-Video	2022	Generate video from text input
34	Midjourney [191]	Midjourney, Inc	Text-to-Image	2022	Generate realistic and creative image from text prompt
35	Perplexity [192]	Perplexity.ai	Text-to-Text	2022	
36	Whisper [193]	OpenAI	Text-to-Speech	2022	Speech recognition and translation
37	chatGPT [1]	OpenAI	Text-to-Text	2022	Conversational chatbot that generates human-like text responses
38	Stable diffusion [194]	Stability AI	Text-to-Image	2022	Generate photo-realistic images given any text input

TABLE 3. *(Continued.)* GAI tools.

39	Dreamstudio [195]	Stability AI	Text-to-Image	2022	Generate photo-realistic images given any text input
40	ChatSonic [196]	Writesonic	Text-to-Image, Text-to-Text,	2022	Conversational chatbot which can generate human text response and image
41	Firefly [197]	Adobe	Text-to-Image	2023	Generative image from text prompt
42	Wordtune Spice [198]	AI21 Labs	Text-to-Text	2023	Writing Generator
43	Altered Studio [199]	Altered	Text-to-Speech	2023	Voice generator
44	Amper [200]	Amper	Text-to-Music	2023	Generate Music of various genres
45	Cloude 2 [201]	Anthropic	Text-to-Text	2023	Content Generation, AI Assistant
46	appleGPT [202]	Apple	Text-to-Text	2023	chatbot summarize text and answer questions
47	Pi [203]	inflection	Text-to-Text	2023	
48	Canva AI [204]	Canva	Text-to-Image	2023	Generate image from text input
49	PaLM 2 [205]	Google	Text-to-Text	2023	Generate code, creative content, Translation and Q&A
50	Codey [206]	Google	Text-to-Code	2023	Generate Code based on user input
51	Imagen [147]	Google	Text-to-Image	2023	Generate realistic image
52	Bard (Gemini) [207]	Google	Text-to-Text	2023	Conversational chatbot that generates human-like text responses
53	Phenaki [208]	Google	Text-to-Video	2023	Generate video from text input of animate objects
54	parti [209]	Google	Text-to-Image	2023	Can Generate realistic image from text prompt
55	MusicLM [210]	Google	Text-to-Music	2023	Generate Music of various genres
56	Studio bot [211]	Google	Text-to-Code	2023	Code Companion for android developer
57	StarCoder [150]	Hugginface + ServiceNow	Text-to-Code	2023	state-of-the-art large language model (LLM) for code
58	VoiceBox [212]	Meta AI	Text-to-Speech	2023	Generate voice clips
59	Metamate [213]	Meta AI	Text-to-Code	2023	Software debugging
60	Shepherd [214]	Meta AI	Text-to-Text	2023	Improve the accuracy of AI generated response
61	CM3leon [215]	Meta AI	Text-to-Image	2023	generate text and images
62	AudioCraft [216]	Meta AI	Text-to-Music	2023	Music Generator
63	MusicGen [217]	Meta AI	Text-to-Music	2023	Music Generator
64	DALL-E 2 [218]	OpenAI	Text-to-Image	2023	Generate image from text description
65	chatGPT plus (GPT-4) [105]	OpenAI	Text-to-Text	2023	Advanced ChatGPT, Conversational chatbot that generates human-like text responses , it's a paid version.
66	RoomGPT [219]	RoomGPT.io	Text-to-Image	2023	Design home and room
67	Gen-2 [220]	RunwayML	Text-to-Video	2023	Design video from text input
68	Einstein GPT [221]	SalesForce	Text-to-Text	2023	Chatbot built in top for chatGPT which generate text, translate languages, write different kinds of creative content, and answer questions
69	Scribe [222]	Scribe AI	Text-to-Text	2023	Creates Documentation, how-to guides, SOPs and training manuals
70	Devin [223]	Cognition Labs	Text-to-Code	2024	Can write the entire app,
71	Sora [224]	OpenAI	Text-to-Video	2024	Create realistic video scene, safety step testing is underway.
72	Speechelo [225]	speechelo	Text-to-Speech	N/A	Creates realistic voice from text
73	Kits [226]	kits AI	Text-to-Speech	N/A	Voice generator
74	Voice Over [227]	Speechify	Text-to-Speech	N/A	Creates natural Voiceovers for any Content
75	quillbot [228]	Course Hero	Text-to-Text	N/A	Can paraphrase, rewrite the text
76	WellSaid [229]	WellSaid Lab	Text-to-Speech	N/A	Voice generator
77	Jeda [230]	Jeda.Ai	Text-to-Image	N/A	Turns idea to flowchart

On the other hand, reverse process operates oppositely by progressively denoising the noisy data starting from the random sample $X_t \sim N(0, I)$ toward generating the original distribution (Posterior Distribution) defined on equation 13.

$$P_\theta\left(X_{t-1} \mid X_t\right)=N\left(X_{t-1} ; \mu_\theta\left(X_t, t\right), \sum\nolimits_\theta\left(X_t, t\right)\right) \tag{13}$$

Markov chain (see equation 14) is applied to maintain consistency during denoising process.

$$P_\theta\left(X_0: T\right)=P\left(X_t\right) \prod_{t=1}^T P_\theta\left(X_{t-1} \mid X_t\right) \tag{14}$$

The log likelihood is employed to optimize the learning process towards the original data distribution, which in turn leads to an increase in the variational lower bound. To minimize the negative variational lower bound, the Kullback-Leibler (KL) divergence is utilized as depicted on equation 15

$$\begin{aligned}\mathcal{L}_{vlb}= & -\log_{P_\theta}\left(X_0 \mid X_1\right)+KL\left(P\left(X_T \mid X_0\right) || \pi\left(X_T\right)\right) \\ & +\sum_{t>1} KL(p\left(X_t \mid X_{t-1}\right), X_0) || P_\theta\left(X_{t-1} \mid X_t\right))\end{aligned} \tag{15}$$

whereby L_{vlb} infers log likelihood variation lower bound.

The study by Ho et al. [135] simplified $\mathcal{L}_{vlb}$ by removing weighting coefficient resulting $\mathcal{L}_{simple}$ defined by equation 16.

$$L_{Simple}=E_{t, X_0, \epsilon}[||\epsilon-\epsilon_\theta(X_t, t)||^2]. \tag{16}$$

1) SCORE-BASED GENERATIVE MODELS (SGMS)

The formulation of **SGMs** relies on the concept of a score (Stein) function, defined as the gradient of the logarithm of the probability density $\nabla_x \log P(x)$ [137] This approach involves perturbing data with Gaussian noise progressively and employing a noise-conditional score model (NCSN), which is a neural network model to estimate the score function [138] burden. Mathematically is derived as $\nabla_{X_t} \log P_{\sigma_t}\left(X_t \mid X\right)=-\frac{X_t-X}{\sigma_t}$ with noise distribution of whereby noise distribution $P_{\sigma_t}\left(X_t \mid X\right)=N(X_t, X, \sigma_t^2 I)$ and σ_t is a sequence of noise level.

2) STOCHASTIC DIFFERENTIAL EQUATIONS (SDE)

SDEs formulation is continuous diffusion process that can generalize the prior mentioned formation DDPMs and SGMs perturbation and denoising process [139].

$$dx=f(x, t)\, dt+g(t)\, dw \tag{17}$$

Given that Equation 17 defines forward SDE which estimate the score function whereby f is the function of x, and t for drift coefficient g(t) is the diffusion coefficient, dt denotes infinitesimal negative time step and w infer standard wiener process. To denoise the date the reverse process is required on the forward SDE (Equation 19). It is given as:

$$dx=\left[f(x, t)-g^2(t) \nabla_x \log p_t(x)\right] dt+g(t)\, d\overline{w} \tag{18}$$

where $\overline{w}$ represent Brownian motion. Reverse SDE [140] can be solved numerically using a trained neural network $S_\theta(x, t) \approx \nabla_x \log p_t(x)$ which compute the actual score function using the objective depicted in equation 19:

$$\begin{aligned}\mathcal{L}= & \mathrm{E}_t[(\lambda(t)\, \mathrm{E}_{p(x_0)} \mathrm{E}_{P_t}(x_t \mid x_0) || S_\theta(x, t) \\ & -\nabla_x \log p_t(x_t \mid x_0) ||_2^2]\end{aligned} \tag{19}$$

where λ is the weighted function and sampling can be accomplished by using numerical method such as Euler-Maruyama, Prediction-Correction, prediction flow ODE.

III. GAI TASK AND TOOLS

Task: GAI encompasses a wide array of tasks; these tasks involve the generation on new data from the given input. including Speech Generation (Text-to-Speech), Image Generation (Text-to-Image), Text Generation (Text-to-Text), Code Generation (Text-to-Code), Music Generation (Text-to-Music), Video Generation (Text-to-Video), and Scientific Content Generation (Text-to-Science). These tasks are supported by various cutting-edge tools. Below, we explore these tasks, providing accompanying examples of tool outputs in various use cases.

Tools: As depicted in Table 3, numerous tools have been developed by various companies to address a variety of real-world problems, as mentioned earlier. Since the emergence of ChatGPT, there has been an exponential increase in the release of GIA tools. The usage and demand for these tools have surged significantly. They are no longer limited to research purposes but are now being utilized on a daily basis and in commercial applications. Individuals from various roles and expertise levels are eager to adopt these tools, as highlighted by the report from Deloitte [141].

Referring to Table 3, we present a comprehensive discussion of the most well-known and widely used tools to the best of our knowledge. We include details such as their functionality, developer, year of release, and categorization. Most of these tools were unveiled in 2023, though a few were introduced earlier.In certain instances, release dates may be labeled as N/A due to unavailability of information. Notably, Google emerges as the primary developer of a multitude of generative tools, closely followed by Meta AI and OpenAI.

Input: In many GAI tasks, the primary input is text, commonly referred to as a prompt. This prompt is crucial in determining the generated output, making **prompt engineering** a vital skill [142]. Prompt engineering involves designing inputs for AI tools to produce optimal outputs. A well-crafted prompt may include instructions, context, output indicators, and input data [143]. Clear instructions within the prompt may entail providing detailed query information, utilizing delimiters to separate different parts of the input, breaking down tasks into subtasks, offering examples or references, and specifying the desired length and format of the output [144]. These components ensure that the AI model understands the task at hand and effectively generates outputs meeting the desired criteria.

A. TEXT GENERATION

This task involves taking text as input and generating corresponding text-based responses. It is often associated with question-and-answer conversational systems, commonly called chatbots. Many renowned GAI tools fall within this category, with ChatGPT being a groundbreaking example in the field of GAI. Other notable tools in this category include Google's Bard, OpenAI's ChatGPT Plus, Wordtune Spice, and Cohere's Generate. We conducted a comprehensive performance assessment of two prominent and renowned text-to-text tools, Bard and ChatGPT. Both were presented with identical queries: '*Provide a brief description of what Bard is in one paragraph*', '*Provide a brief description of what ChatGPT is in one paragraph*', and a Swahili question, '*Habari za saa hizi*'. The results as illustrated by Figure 9, unmistakably indicate that ChatGPT outperformed Bard in delivering more precise answers to the questions.

B. IMAGE GENERATION

It's a task which encompasses the process of utilizing textual prompts or visual to generate corresponding images, spanning various visual domains, including graphics, photographs, and artwork. As an illustration of text-to-image concept, we conducted experiments using 'Firefly' from Adobe and 'Stable Diffusion' by Stability as our subjects. By prompting these models with '*College Student Programming*', we obtained their respective outputs, as showcased in Figure 10, the results clearly indicate that while 'Firefly' excelled in delivering more precise outputs in alignment with the input, Stable Diffusion exhibited superior image resolution compared to its counterpart. Another scenario image generation revolves around the transformation of an image from one form to another, guided by textual descriptions provided as input. Within this domain, numerous tools have demonstrated promising capabilities in effecting such transformations. Notably, we have explored the performance of RoomGPT and Runaway, as exemplified in Figure 11 and Figure 12, respectively.

C. VIDEO GENERATION

This task involves generating new videos based on textual or visual inputs, whereby visual encompasses a diverse range of content that includes both images and videos. In this domain, there are notable tools designed to accommodate exclusively text-based descriptions as inputs. A prime example is 'Parti' by Google, and DALL E-2 [145] by openAI are proficient tools focused on creating videos solely from textual prompts. Nonetheless, the field of video generation is in a state of continuous evolution. Tools such as 'Gen-2' by RunwayML, 'Imagen Video' by Google [146], and 'Make-A-Video' by Meta [147] have emerged as pioneers. These advanced platforms possess the remarkable capability unlimited to textual descriptions but also seamlessly integrate images and videos as input, transcending conventional boundaries. Their excellence lies in their adeptness at transforming these inputs into entirely novel video compositions, thus unveiling the exciting potential of GAI in the creative realm of video production.

D. CODE GENERATION

Code generation tools are specialized software utilities capable of automatically producing code blocks for various programming languages based on textual descriptions provided as input [148]. These tools leverage sophisticated models trained on extensive publicly available code repositories, boasting billions of parameters. Their primary objective is to assist human developers by comprehending plain English and translating it into functional code. Notable examples of such tools include StarCoder [149], Codex [150], CoPilot, Codey, and Code Interpreter. Additionally, it's worth noting that several text-to-text tools, including ChatGPT and Bard as depicted by Figure 13, also possess the capacity to generate code.

E. MUSIC GENERATION

It's a fascinating generative task involving entirely new music's composition. This innovative process takes input in various forms, including textual descriptions, sequences of musical notes, and even audio samples [151]. The objective is to harness these inputs and transform them into fresh musical compositions that encapsulate rhythm, melody, harmonious chords, and diverse musical instruments. Prominent tools like MuseNet [152] and Jukebox [153] stand out as prime examples in the music generation. These innovative platforms harness the power of GAI to craft musical compositions spanning various genres and styles. They excel in infusing creativity into the art of music, opening new avenues for artists and enthusiasts to explore and enjoy.

F. SPEECH GENERATION

The generation of human-like speech or voice relies on textual or audio input [155]. Textual input can encompass written text, such as sentences, paragraphs, or entire documents, and it can span multiple languages, including punctuation, special symbols, and formatting instructions [156], [157]. Speech generation models undertake a sequence of steps that involve speech synthesis, enhancement, and conversion. The enhancement process includes noise handling, tone modulation, emotion conveyance, and other nuanced features [158], [159]. Numerous tools have been developed in this domain to facilitate speech generation, some of which include Whisper, Speechelo, Synthesys, Voice Over, and WaveNet. These tools are proficient in generating voices or speech that closely mimic natural language, effectively blurring the line between human and artificial speech synthesis.

G. SCIENTIFIC CONTENT GENERATION

Scientific content generation is a multifaceted process encompassing the creation of informative and scholarly content across various domains of science, including mathematics, physics, chemistry, and biology. This endeavor seeks to harness the power of GAI to produce content that is accurate and insightful, aiding in disseminating scientific knowledge. One notable study in this field, conducted by Rodriguez et al. [160], delved into the innovative way of generating scientific figures based on textual input. This groundbreaking research leveraged diffusion models to seamlessly translate textual descriptions into visually informative scientific figures, thereby streamlining the process of scientific communication and visualization. Furthermore, Google's ongoing research project, Minerva [161], represents a significant stride in solving quantitative reasoning problems. This initiative harnesses the capabilities of Large Language Models (LLMs) to tackle complex quantitative challenges, thereby enhancing our understanding of mathematics and its practical applications within the scientific landscape. In parallel, Galactica [162], a cutting-edge tool developed by Meta AI, plays a pivotal role in scientific writing. This platform equips scientists and researchers with powerful tools to streamline articulating their scientific discoveries, theories, and insights.

IV. INDUSTRIAL APPLICATION OF GIA

GAI technology's relevance in the present and future is indispensable. Currently, GAI is exerting an exponential impact across a broad spectrum of industries, and this section will delve into a detailed exploration of the sectors that are most impacted.

A. MEDIA AND ENTERTAINMENT

In the entertainment industry, GAI models are beginning to have a significant impact despite being in their early stages. Their influence spans various entertainment domains, encompassing scriptwriting and storyboarding for novels, plays, and films, audio production [229] involving composition, arrangement, and mixing, game design and character creation, the creation of captivating virtual worlds, marketing campaigns, and the generation of both moving and static images. Notably, a wide range of accessible tools, as demonstrated in Table 3, make it easier to generate content such as reels, jokes, and images [230]. Many of these tools are cost-effective or even free, providing an alternative to traditional content creation methods. As an illustration of their potential, in 2022, RunwayAI played a role in creating the Academy Award-winning film ''Everything Everywhere All at Once'' which received recognition with seven Oscars awards [231], [232].

B. EDUCATION AND RESEARCH

GAI is rapidly reshaping the educational landscape, offering innovative solutions that elevate the learning experience for both students and educators. One significant impact of GAI in education is the emergence of personalized content generation tools. Exemplified by technologies like GPT-3, GPT-4 and Bard, these tools empower educators to craft tailored learning materials, including interactive lessons, quizzes, and study guides, precisely catering to the unique needs of individual students and instructors [233]. Furthermore, AI-driven chatbots and virtual tutors provide students with real-time support, offering explanations, addressing queries, and delivering personalized feedback [35]. This transformative technology holds the potential to reinvent how students access and engage with educational content, promoting accessibility and adaptability according to each learner's specific preferences [234], [235].

GAI has also opened new avenues of research and academic exploration. The rapid development of GAI tools has piqued the interest of researchers and academics across the globe, leading to an array of research opportunities [236]. Tech giants and research institutions are investing significant resources to explore and invent new tools and technologies in this field. This is evident in the surge of publications related to GAI, both in peer-reviewed databases like IEEE and non-reviewed platforms like arXiv, where GAI topics have gained prominence. The fusion of education and GAI has not only transformed the learning experience but has also sparked a thriving academic domain that promises continued growth and innovation [237].

C. HEALTHCARE

GAI is making substantial inroads in healthcare, particularly in medical imaging [238]. It plays a crucial role in overcoming challenges related to limited datasets by enabling the synthesis of new data [239], [240], ultimately enhancing the quality and diversity of medical images. This innovation is set to revolutionize disease detection and diagnosis, providing healthcare professionals with more accurate and detailed information. In addition, GAI is transforming the administrative aspects of patient care. Streamlining administrative processes and offering virtual health assistants simplifies healthcare management and provides personalized health advice, medication reminders, and emotional support [241]. Moreover, GAI is revolutionizing treatment planning. Leveraging patient-specific data, it can generate customized treatment plans tailored to an individual's genetic makeup, lifestyle, and medical history. This approach represents a significant leap toward precision medicine, ensuring patients receive the most effective and personalized treatment.

Furthermore, GAI is playing a pivotal role in the realm of drug development and discovery [242], [243], [244]. Through the generation of molecular structures [245] and predictive modeling, it expedites the identification of novel therapeutic compounds. These advancements can address previously untreatable diseases, instilling hope in countless patients across the globe. Notably, the collaboration between NVIDIA and Evozyne in implementing GAI, specifically ProT-VAE, signifies the remarkable synergy between AI and the healthcare sector. By employing the Protein Transformer Variational AutoEncoder, they have laid the groundwork for creating synthetic proteins [246], opening up new avenues for therapeutic solutions in the fight against challenging incurable diseases. Yet another noteworthy example is the collaborative research venture between Google and Cognizant [247]. Their joint effort aims to construct a Large Language Model (LLM) tailored for healthcare applications, specifically focusing on enhancing Healthcare administrative tasks. This endeavor harnesses the capabilities of Google Cloud and its framework to create cutting-edge GAI solutions for the healthcare sector.

D. BUSINESS

GAI has firmly established its presence in the business landscape. Many of the applications listed in Table 3 operate on a subscription-based model, reflecting the growing commercial nature of these tools. Bloomberg Intelligence predicts that GAI will generate 137 billion US dollars in 2023 and is expected to surge to 1.3 trillion US dollars by 2030 [248]. This profound impact extends across various industries, from manufacturing and wholesale to retail businesses, banking, agriculture, and many more. GAI's reach spans from creating new products and automating financial data analysis to generating personalized advertising campaigns [249], [250], [251], offering tailored product recommendations to customers, and producing product descriptions and news articles [252]. It is increasingly evident that GAI is reshaping the business landscape and holds immense economic potential in the future.

For example, Amazon is actively harnessing GAI capabilities to empower sellers in crafting engaging, compelling, and effective product listings through brief descriptions of their products. Amazon leverages GAI to generate high-quality content, which sellers can further refine or directly submit to enrich the Amazon catalog [253].

Improved Business Network Security performance, GAI significantly can impact network security by facilitating traffic anomaly prediction. For instance, Synthetic traffic packets generated by GAI algorithms like GANs [254], [255], [256], [257] can simulate real-world scenarios, enabling the testing of network responses to detect anomalies effectively. This technique aids in pinpointing vulnerable loopholes, thus enabling the establishment of robust network mechanisms to deter intrusion attempts [258], [259]. Also, GAI is crucial in analyzing network traffic patterns and devising strategies for optimizing bandwidth allocation, congestion control, and routing. By leveraging GAI's capabilities, networks can achieve enhanced performance and efficiency and affect business performance.

V. CHALLENGES OF GAI

Despite the abundant benefits brought by GAI into our daily life, it has also raised new uncertainties in various areas life, as discussed below:

Job deterioration; optimizing and automating business processes are anticipated to replace many existing careers with creative and GAI functions. GAI's impact on the labor market is poised to transform the employment landscape, gradually replacing many traditional roles with advanced technology. According to the World Economic Forum's report [260], tasks with the highest potential for automation by Large Language Models (LLMs) are routine and repetitive. These tasks include those performed by Credit Authorizers, Checkers, Clerks, Management Analysts, Telemarketers, Statistical Assistants, and Tellers [261], [262]. Therefore, individuals must prioritize reskilling and adaptability to prepare for AI-driven jobs in the future effectively.

Privacy and Security concern; The cybersecurity infrastructure domain is presently undergoing a profound and rapid transformation, primarily driven by the integration of GAI. This substantial shift is giving rise to a host of pressing concerns and challenges for the future like: **Sophisticated cyberwarfare**, currently, we are witnessing a notable surge in malicious activities, and this trend is expected to continue its upward trajectory while also becoming more intricate and sophisticated [263]. For instance, the emergence of cutting-edge cyber threat tools like WormGPT and FraudGPT [264], [265], which have rapidly established themselves as pioneering elements in cyber threats often referred to as "exclusive bots" [266] by their perpetrators, are engineered to be highly sophisticated and evasive. Moreover, the emergence of increasingly automated and sophisticated malware and ransomware, powered by GAI [267], presents a menacing potential for subverting existing encryption methods [268]. This is primarily due to the immense computational prowess inherent in GAI. As these malicious entities persist and advance, they represent a formidable challenge to the cybersecurity landscape, testing the limits of the resilience and robustness of contemporary cybersecurity systems and protocols [269]. The consequences of these developments are far-reaching, with the prospect of malicious AI proving to be devastating to a nation's critical infrastructure, particularly in scenarios involving state-sponsored or malevolent cyber terrorism [270].

Networks can be easily attached by hackers whereby adversaries could exploit GAI-generated traffic to penetrate networks, as such traffic may appear genuine and evade detection mechanisms. Therefore, while GAI offers avenues for bolstering network security and performance, vigilance against potential misuse is essential.

Increased impersonation and misinformation, escalation of AI advancements across various domains, visual, speech, audio, and text-based applications, has significantly elevated concerns surrounding personal privacy breaches and impersonation. A pertinent example is the music industry, where AI-driven ghostwriters have released a fake audio track emulating the voices of renowned artists like Drake and The Weekend, both of whom are global music sensations [271]. Tracks like "Heart on My Sleeve" and "Cuff It," featuring AI-rendered versions of Rihanna and Beyoncé's voices [272], have garnered attention for their remarkably convincing mimicry. Consequently, the creative industry faces substantial threats, particularly sectors reliant on advanced artificial intelligence. As reported, these technologies can potentially jeopardize careers within the entertainment industry.

Incorrect and hallucinations answers; it is crucial for everyone utilizing GIA tools to acknowledge the possibility of generated answers being incorrect, outdated, or even hallucinated. For instance, in a case study involving Figure 8(d) bard (now Gemini), the prompt "Habari za saa hizi," a Swahili phrase meaning "How are you," had a response of breaking news accompanied with hyperlinks, which was inaccurate and potentially hallucinated. Erroneous responses have been observed in advanced prompts or questions, such as coding-related ones. Therefore, it is imperative to understand that complete reliance on GIA-generated responses is unwarranted. Thorough assessment and review of responses are necessary before applying them to real-world problems. This ensures the validity and reliability of the information generated by GIA and mitigates the risks associated with potential inaccuracies or hallucinations.

VI. THE FUTURE OF GAI

GAI undoubtedly holds a significant and promising future, offering a plethora of tangible and transformative possibilities across various domains. However, it is equally accompanied by considerable uncertainty and a range of concerns that deserve in-depth exploration. This section explores the multifaceted aspects of GAI, addressing its potential and the challenges and uncertainties ahead.

A. PIONEER OF FIFTH INDUSTRIAL REVOLUTION (5IR)

GAI represents the promising frontier of the fifth industrial revolution (5IR), a force poised to revolutionize the fourth industrial revolution and create transformative changes across various sectors. This transformation is made possible by the profound interconnection of internet infrastructure, extensive datasets, and distributed computing resources that transcend geographical boundaries. Several industries, including Healthcare, Security, Cyber Infrastructure, Entertainment, and Education, are on the verge of significant disruption due to GAI's capabilities. However, it's crucial to recognize that this disruptive potential may also bring about infrastructure reforms across multiple sectors, potentially leading to high levels of automation and optimization in various career fields.

On Healthcare Industry, as we have witnessed, GAI is already playing a pivotal role in drug discovery, with a particular emphasis on exploring protein molecules. The potential for this technology in the field of drug development is vast, and substantial investments from major technology companies underscore the anticipated advancements in the near future. However, the impact of GAI extends far beyond drug development, as it is expected to transform the patient experience within the healthcare sector fundamentally. By harnessing patients' medical history data, it can autonomously diagnose medical conditions by analyzing metadata like age, sex, and underlying medical conditions. Moreover, it can sift through extensive patient data to identify patterns, make predictions, and suggest appropriate medications. This transformation is set to prioritize patient-centered clinical experiences and drive cost-effectiveness, ultimately leading to significant enhancements in healthcare protocols [273].

Enhanced Entertainment, In the foreseeable future, we stand at the threshold of a transformative era where GAI will likely dominate the realm of content creation in entertainment and media. From crafting intricate scripts and narratives to meticulously arranging scenes and bringing characters to life, the influence of GAI is set to permeate every facet of content generation in these industries. Furthermore, the potential impact is so profound that it might even challenge the boundaries of life and art. Deceased artists could potentially continue to release new albums and creative works, effectively transcending the limitations of mortality. Not only will this innovation usher in a new age of artistic exploration, but it also promises significant cost savings, revolutionizing the economics of movie and music production. Automating scene creation and content generation will reduce expenses and make the creation process more efficient.

New learning era, the advent of AI chatbots like ChatGPT and Google Bard, along with other innovative tools, serves as compelling evidence of the democratization of GAI in the education and learning endeavors. This remarkable progress has rendered the current educational system and resources outdated, particularly in developed countries. It anticipates a comprehensive overhaul of the education system, including teaching resources, to adapt to the exponential growth in the GAI era, aiming to provide highly personalized and adaptive learning experiences.

Advanced Manufacturing Industries, before the emergence of GAI, robotics had already showcased impressive capabilities. However, with the integration of GAI, we can look forward to truly remarkable advancements. Just envision the consequences of infusing GAI into military technology, where we might see the development of generative nuclear weaponry, the formulation of chemical recipes for beverages, detergents, and various industrial products, and the widespread adoption of self-driving vehicles. The range of possibilities is extensive, and it undoubtedly signifies the onset of a new era—an industrial revolution that promises a thoroughly transformed landscape and innovative approaches across numerous sectors of industries.

B. JOB MARKET SHIFTING

The influence of GAI on the labor market is two-fold:

Firstly, it ushers in **new employment opportunitie**s in emerging domains such as AI Explainability and GAI engineering. McKinsey's analysis [274] suggests a gradual rise in job openings within professions exposed to GAI, On and this trend is expected to persist until roughly 2030. A noteworthy revelation is that a substantial 84% of the U.S. workforce occupies positions with the potential to leverage GAI for automating a significant portion of repetitive tasks, leading to a considerable surge in overall productivity. Significantly, 47% of U.S. executives express confidence that integrating GAI will lead to heightened productivity across diverse industries [275], [276].

VII. CONCLUSION

Throughout this paper, we have delved into state-of-the-art models, explored their mathematical foundations, scrutinized their architectural intricacies, and anticipated their evolution in the future. We have also examined prominent tasks, benchmarked state-of-the-art tools against GAI, and assessed their real-world applications. The realms of impact, challenges, and future prospects of GAI have been thoroughly addressed. Indeed, GAI opens the door to a new world filled with both unprecedented numerous opportunities and inherent risks. GIA is a new paradigm that is increasingly influencing various facets of life, including education and business. There is a pressing need to educate the public about the current and future impact of GIA across different domains.

Given the unstoppable emergence and exponential growth of GIA, which is anticipated to play a pivotal role in the advent of the 5IR and the transformation of job markets, as well as revolutionizing cybersecurity practices, it has become apparent that embracing GIA is essential for any business or organization to thrive in this era, therefore it is imperative for managerial teams across various sectors, such as educational institutions, to strategically adopt this technology shift. Establishing compliance policies and risk management plans for managing the impact of GIA is strongly advised. Moreover, prioritizing talent and skill development, along with continuous learning initiatives on GIA, is essential for all workers and the entire organization [277]. This approach ensures keeping track of new developments and effectively leveraging the advantages of GIA, while also being prepared to handle any potential threats it may pose.

Many unanswered questions remain, prompting future research in GIA. Researchers must actively engage in studying and innovating solutions to the challenges posed by GIA such as misinformation, inaccurate responses, cyber threats and job deterioration. For instance, there is a need for research into detecting content generated by GIA tools, ensuring authenticity and originality. Another vital area for research involves the study of explainability and privacy preservation regarding the output of GIA, prioritizing transparency and safeguarding user data Additionally, researchers are expected to devise solutions to combat cyber warfare originating from GIA tools, addressing issues such as impersonation and misinformation. Moreover, research is needed to enhance the confidence, reliability, and timeliness of generated responses. Generally, these avenues offer promising opportunities for continued exploration and advancement within GAI.

REFERENCES

[1] *Introducing ChatGPT*. Accessed: Sep. 30, 2023. [Online]. Available: https://openai.com/blog/chatgpt
[2] C. Leiter, R. Zhang, Y. Chen, J. Belouadi, D. Larionov, V. Fresen, and S. Eger, "ChatGPT: A meta-analysis after 2.5 months," 2023, *arXiv:2302.13795*.
[3] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the Dartmouth summer research project on artificial intelligence," *AI Mag.*, vol. 27, no. 4, pp. 1–12, 2006.
[4] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *J. Ind. Inf. Integr.*, vol. 23, Sep. 2021, Art. no. 100224, doi: 10.1016/j.jii.2021.100224.
[5] N. J. Nilsson, *The Quest for Artificial Intelligence*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[6] P. Hamet and J. Tremblay, "Artificial intelligence in medicine," *Metabolism*, vol. 69, pp. S36–S40, Apr. 2017, doi: 10.1016/j.metabol.2017.01.011.
[7] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach*. Springer, 2013.
[8] D. Li and D. Yi, *Artificial Intelligence With Uncertainty*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2016, doi: 10.1201/9781315366951.
[9] S. Bengesi, T. Oladunni, R. Olusegun, and H. Audu, "A machine learning-sentiment analysis on monkeypox outbreak: An extensive dataset to show the polarity of public opinion from Twitter tweets," *IEEE Access*, vol. 11, pp. 11811–11826, 2023, doi: 10.1109/ACCESS.2023.3242290.
[10] R. Olusegun, T. Oladunni, H. Audu, Y. Houkpati, and S. Bengesi, "Text mining and emotion classification on monkeypox Twitter dataset: A deep learning-natural language processing (NLP) approach," *IEEE Access*, vol. 11, pp. 49882–49894, 2023, doi: 10.1109/ACCESS.2023.3277868.
[11] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958, doi: 10.1037/h0042519.
[12] J. N. Morgan and J. A. Sonquist, "Problems in the analysis of survey data, and a proposal," *J. Amer. Stat. Assoc.*, vol. 58, no. 302, pp. 415–434, Jun. 1963, doi: 10.1080/01621459.1963.10500855.
[13] E. Fix and J. L. Hodges, "Discriminatory analysis. Nonparametric discrimination: Consistency properties," *Int. Stat. Rev. Revue Internationale de Statistique*, vol. 57, no. 3, p. 238, Dec. 1989, doi: 10.2307/1403797.
[14] J. Platt. (Apr. 1998). *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines*. Accessed: Sep. 30, 2023. [Online]. Available: https://www.microsoft.com/en-us/research/publication/sequential-minimal-optimization-a-fast-algorithm-for-training-support-vector-machines/
[15] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/a:1010933404324.
[16] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
[17] J. J. Hopfield and D. W. Tank, "'Neural' computation of decisions in optimization problems," *Biol. Cybern.*, vol. 52, no. 3, pp. 141–152, Jul. 1985, doi: 10.1007/bf00339943.
[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
[19] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997, doi: 10.1109/78.650093.
[20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
[21] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.
[22] S. Cong and Y. Zhou, "A review of convolutional neural network architectures and their optimizations," *Artif. Intell. Rev.*, vol. 56, no. 3, pp. 1905–1969, Mar. 2023, doi: 10.1007/s10462-022-10213-5.
[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2012. Accessed: Sep. 30, 2023. [Online]. Available: https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html
[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778. Accessed: Sep. 30, 2023. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

[25] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269. Accessed: Sep. 30, 2023. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html

[26] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[27] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Mach. Learn.*, May 2019, pp. 6105–6114. Accessed: Sep. 30, 2023. [Online]. Available: https://proceedings.mlr.press/v97/tan19a.html

[28] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017. Accessed: Aug. 15, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[29] I. J. Goodfellow et al., "Generative adversarial networks," 2014, *arXiv:1406.2661*.

[30] C. Zheng, G. Wu, F. Bao, Y. Cao, C. Li, and J. Zhu, "Revisiting discriminative vs. generative classifiers: Theory and implications," 2023, *arXiv:2302.02334*.

[31] E. Brophy, Z. Wang, Q. She, and T. Ward, "Generative adversarial networks in time series: A systematic literature review," *ACM Comput. Surv.*, vol. 55, no. 10, pp. 1–31, Feb. 2023, doi: 10.1145/3559540.

[32] G. Zhou, S. Xie, G. Hao, S. Chen, B. Huang, X. Xu, C. Wang, L. Zhu, L. Yao, and K. Zhang, "Emerging synergies in causality and deep generative models: A survey," 2023, *arXiv:2301.12351*.

[33] N. R. Mannuru, S. Shahriar, Z. A. Teel, T. Wang, B. D. Lund, S. Tijani, C. O. Pohboon, D. Agbaji, J. Alhassan, J. Galley, R. Kousari, L. Ogbadu-Oladapo, S. K. Saurav, A. Srivastava, S. P. Tummuru, S. Uppala, and P. Vaidya, "Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development," *Inf. Develop.*, Sep. 2023, Art. no. 02666669231200628, doi: 10.1177/02666669231200628.

[34] E. Brynjolfsson, D. Li, and L. R. Raymond, "Generative AI at work," Nat. Bur. Econ. Res., Cambridge, MA, USA, Working Paper, 2023, doi: 10.3386/w31161.

[35] J. Qadir, "Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, May 2023, pp. 1–9, doi: 10.1109/EDUCON54358.2023.10125121.

[36] B. Meskó and E. J. Topol, "The imperative for regulatory oversight of large language models (or generative AI) in healthcare," *NPJ Digit. Med.*, vol. 6, no. 1, pp. 1–6, Jul. 2023, doi: 10.1038/s41746-023-00873-0.

[37] S. Feuerriegel, J. Hartmann, C. Janiesch, and P. Zschech, "Generative AI," *Bus. Inf. Syst. Eng.*, vol. 66, no. 1, pp. 111–126, Feb. 2024, doi: 10.1007/s12599-023-00834-7.

[38] W. Liao, X. Lu, Y. Fei, Y. Gu, and Y. Huang, "Generative AI design for building structures," *Autom. Construction*, vol. 157, Jan. 2024, Art. no. 105187, doi: 10.1016/j.autcon.2023.105187.

[39] M. Xu, H. Du, D. Niyato, J. Kang, Z. Xiong, S. Mao, Z. Han, A. Jamalipour, D. I. Kim, X. Shen, V. C. M. Leung, and H. V. Poor, "Unleashing the power of edge-cloud generative AI in mobile networks: A survey of AIGC services," *IEEE Commun. Surveys Tuts.*, 2024, doi: 10.1109/COMST.2024.3353265.

[40] A. Karapantelakis, P. Alizadeh, A. Alabassi, K. Dey, and A. Nikou, "Generative AI in mobile networks: A survey," *Ann. Telecommun.*, vol. 79, nos. 1–2, pp. 15–33, Feb. 2024, doi: 10.1007/s12243-023-00980-9.

[41] M. Zhou, V. Abhishek, T. Derdenger, J. Kim, and K. Srinivasan, "Bias in generative AI," 2024, *arXiv:2403.02726*.

[42] S. C. Matz, J. D. Teeny, S. S. Vaid, H. Peters, G. M. Harari, and M. Cerf, "The potential of generative AI for personalized persuasion at scale," *Sci. Rep.*, vol. 14, no. 1, p. 4692, Feb. 2024, doi: 10.1038/s41598-024-53755-0.

[43] Y. K. Dwivedi, N. Pandey, W. Currie, and A. Micu, "Leveraging ChatGPT and other generative artificial intelligence (AI)-based applications in the hospitality and tourism industry: Practices, challenges and research agenda," *Int. J. Contemp. Hospitality Manage.*, vol. 36, no. 1, pp. 1–12, Jan. 2024, doi: 10.1108/ijchm-05-2023-0686.

[44] T. K. F. Chiu, "Future research recommendations for transforming higher education with generative AI," *Comput. Educ., Artif. Intell.*, vol. 6, Jun. 2024, Art. no. 100197, doi: 10.1016/j.caeai.2023.100197.

[45] A. Bandi, P. V. S. R. Adapa, and Y. E. V. P. K. Kuchi, "The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges," *Future Internet*, vol. 15, no. 8, p. 260, Jul. 2023, doi: 10.3390/fi15080260.

[46] H. Al Naqbi, Z. Bahroun, and V. Ahmed, "Enhancing work productivity through generative artificial intelligence: A comprehensive literature review," *Sustainability*, vol. 16, no. 3, p. 1166, Jan. 2024, doi: 10.3390/su16031166.

[47] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," 2021, *arXiv:2003.05991*.

[48] U. Michelucci, "An introduction to autoencoders," 2022, *arXiv:2201.03898*.

[49] J. Zhai, S. Zhang, J. Chen, and Q. He, "Autoencoder and its various variants," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2018, pp. 415–419, doi: 10.1109/SMC.2018.00080.

[50] D. P. Kingma, S. Mohamed, D. J. Rezende, and M. Welling, "Semi-supervised learning with deep generative models," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2014. Accessed: Aug. 7, 2023. [Online]. Available: https://proceedings.neurips.cc/paper/2014/hash/d523773c6b194f37b938d340d5d02232-Abstract.html

[51] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *Found. Trends Mach. Learn.*, vol. 12, no. 4, pp. 307–392, 2019, doi: 10.1561/2200000056.

[52] H. Akrami, A. A. Joshi, J. Li, S. Aydore, and R. M. Leahy, "Brain lesion detection using a robust variational autoencoder and transfer learning," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 786–790, doi: 10.1109/ISBI45749.2020.9098405.

[53] X. Shen, B. Liu, Y. Zhou, J. Zhao, and M. Liu, "Remote sensing image captioning via variational autoencoder and reinforcement learning," *Knowl.-Based Syst.*, vol. 203, Sep. 2020, Art. no. 105920, doi: 10.1016/j.knosys.2020.105920.

[54] G. Zhao and Y. Peng, "Semisupervised SAR image change detection based on a Siamese variational autoencoder," *Inf. Process. Manage.*, vol. 59, no. 1, Jan. 2022, Art. no. 102726, doi: 10.1016/j.ipm.2021.102726.

[55] H. W. L. Mak, R. Han, and H. H. F. Yin, "Application of variational AutoEncoder (VAE) model and image processing approaches in game design," *Sensors*, vol. 23, no. 7, p. 3457, Mar. 2023, doi: 10.3390/s23073457.

[56] M. A. Yílmaz, O. Keless, H. Güven, A. M. Tekalp, J. Malik, and S. Kíranyaz, "Self-organized variational autoencoders (Self-VAE) for learned image compression," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2021, pp. 3732–3736, doi: 10.1109/ICIP42928.2021.9506041.

[57] Z.-S. Liu, W.-C. Siu, and L.-W. Wang, "Variational AutoEncoder for reference based image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*. Nashville, TN, USA: IEEE, Jun. 2021, pp. 516–525, doi: 10.1109/CVPRW53098.2021.00063.

[58] G. Carbajal, J. Richter, and T. Gerkmann, "Guided variational autoencoder for speech enhancement with a supervised classifier," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 681–685, doi: 10.1109/ICASSP39728.2021.9414363.

[59] T. Srikotr and K. Mano, "Sub-band vector quantized variational AutoEncoder for spectral envelope quantization," in *Proc. IEEE Region 10 Conf. (TENCON)*, Oct. 2019, pp. 296–300, doi: 10.1109/TENCON.2019.8929436.

[60] Z. Wang, "Using Gaussian process in clockwork variational autoencoder for video prediction," in *Proc. Int. Conf. Inf. Technol. Res. Innov. (ICITRI)*, Nov. 2022, pp. 6–11, doi: 10.1109/ICITRI56423.2022.9970241.

[61] M. S. Kim, J. P. Yun, S. Lee, and P. Park, "Unsupervised anomaly detection of LM guide using variational autoencoder," in *Proc. 11th Int. Symp. Adv. Topics Electr. Eng. (ATEE)*, Mar. 2019, pp. 1–5, doi: 10.1109/ATEE.2019.8724998.

[62] C. K. Meher, R. Nayak, and U. C. Pati, "Dual stream variational autoencoder for video anomaly detection in single scene videos," in *Proc. 2nd Odisha Int. Conf. Electr. Power Eng., Commun. Comput. Technol. (ODICON)*, Nov. 2022, pp. 1–6, doi: 10.1109/ODICON54453.2022.10010086.

[63] H. Yanagihashi and T. Sudo, "Noise-robust early detection of cooling fan deterioration with a variational autoencoder-based method," in *Proc. 9th Int. Conf. Condition Monitor. Diagnosis (CMD)*, Nov. 2022, pp. 183–188, doi: 10.23919/CMD54214.2022.9991542.

[64] H. Purohit, T. Endo, M. Yamamoto, and Y. Kawaguchi, ''Hierarchical conditional variational autoencoder based acoustic anomaly detection,'' in *Proc. 30th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2022, pp. 274–278, doi: 10.23919/EUSIPCO55093.2022.9909785.
[65] B. Shen, L. Yao, and Z. Ge, ''Nonlinear probabilistic latent variable regression models for soft sensor application: From shallow to deep structure,'' *Control Eng. Pract.*, vol. 94, Jan. 2020, Art. no. 104198, doi: 10.1016/j.conengprac.2019.104198.
[66] B. Shen and Z. Ge, ''Supervised nonlinear dynamic system for soft sensor application aided by variational auto-encoder,'' *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6132–6142, Sep. 2020, doi: 10.1109/TIM.2020.2968162.
[67] W. Xie, J. Wang, C. Xing, S. Guo, M. Guo, and L. Zhu, ''Variational autoencoder bidirectional long and short-term memory neural network soft-sensor model based on batch training strategy,'' *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5325–5334, Aug. 2021, doi: 10.1109/TII.2020.3025204.
[68] R. Xie, N. M. Jan, K. Hao, L. Chen, and B. Huang, ''Supervised variational autoencoders for soft sensor modeling with missing data,'' *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2820–2828, Apr. 2020, doi: 10.1109/TII.2019.2951622.
[69] F. Guo, R. Xie, and B. Huang, ''A deep learning just-in-time modeling approach for soft sensor based on variational autoencoder,'' *Chemometric Intell. Lab. Syst.*, vol. 197, Feb. 2020, Art. no. 103922, doi: 10.1016/j.chemolab.2019.103922.
[70] L. Li, J. Yan, H. Wang, and Y. Jin, ''Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1177–1191, Mar. 2021, doi: 10.1109/TNNLS.2020.2980749.
[71] N. T. N. Anh, T. Q. Khanh, N. Q. Dat, E. Amouroux, and V. K. Solanki, ''Fraud detection via deep neural variational autoencoder oblique random forest,'' in *Proc. IEEE-HYDCON*, Sep. 2020, pp. 1–6, doi: 10.1109/HYDCON48903.2020.9242753.
[72] C. Zhang, C. Zhang, S. Zheng, Y. Qiao, C. Li, M. Zhang, S. K. Dam, C. M. Thwal, Y. L. Tun, L. L. Huy, D. Kim, S.-H. Bae, L.-H. Lee, Y. Yang, H. T. Shen, I. S. Kweon, and C. S. Hong, ''A complete survey on generative AI (AIGC): Is ChatGPT from GPT-4 to GPT-5 all you need?'' 2023, *arXiv:2303.11717*.
[73] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang, A. Xiao, C. Xu, Y. Xu, Z. Yang, Y. Zhang, and D. Tao, ''A survey on vision transformer,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 87–110, Jan. 2023, doi: 10.1109/TPAMI.2022.3152247.
[74] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, ''Transformers in vision: A survey,'' *ACM Comput. Surv.*, vol. 54, no. 10s, pp. 200:1–200:41, Sep. 2022, doi: 10.1145/3505244.
[75] T. Lin, Y. Wang, X. Liu, and X. Qiu, ''A survey of transformers,'' *AI Open*, vol. 3, pp. 111–132, Jan. 2022, doi: 10.1016/j.aiopen.2022.10.001.
[76] M. Zong and B. Krishnamachari, ''A survey on GPT-3,'' 2022, *arXiv:2212.00857*.
[77] B. Ghojogh and A. Ghodsi, ''Attention mechanism, transformers, BERT, and GPT: Tutorial and survey,'' *Open Sci. Framework*, Dec. 2020, doi: 10.31219/osf.io/m6gcn.
[78] *Generative AI: A Game-Changer Society Needs to be Ready for*. Accessed: Aug. 16, 2023. [Online]. Available: https://www.weforum.org/agenda/2023/01/davos23-generative-ai-a-game-changer-industries-and-society-code-developers/
[79] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, ''Improving language understanding by generative pre-training,'' Tech. Rep., 2018.
[80] I. Solaiman et al., ''Release strategies and the social impacts of language models,'' 2019, *arXiv:1908.09203*.
[81] T. B. Brown et al., ''Language models are few-shot learners,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 1877–1901.
[82] X. Zheng, C. Zhang, and P. C. Woodland, ''Adapting GPT, GPT-2 and BERT language models for speech recognition,'' in *Proc. IEEE Autom. Speech Recognit. Understand. Workshop (ASRU)*, Dec. 2021, pp. 162–168, doi: 10.1109/ASRU51503.2021.9688232.
[83] A. Shrivastava, R. Pupale, and P. Singh, ''Enhancing aggression detection using GPT-2 based data balancing technique,'' in *Proc. 5th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2021, pp. 1345–1350, doi: 10.1109/ICICCS51141.2021.9432283.
[84] M. Tamimi, M. Salehi, and S. Najari, ''Deceptive review detection using GAN enhanced by GPT structure and score of reviews,'' in *Proc. 28th Int. Comput. Conf., Comput. Soc. Iran (CSICC)*, Jan. 2023, pp. 1–7, doi: 10.1109/CSICC58665.2023.10105368.
[85] S. Saravanan and K. Sudha, ''GPT-3 powered system for content generation and transformation,'' in *Proc. 5th Int. Conf. Comput. Intell. Commun. Technol. (CCICT)*, Jul. 2022, pp. 514–519, doi: 10.1109/CCiCT56684.2022.00096.
[86] K. H. Manodnya and A. Giri, ''GPT-K: A GPT-based model for generation of text in Kannada,'' in *Proc. IEEE 4th Int. Conf. Cybern., Cognition Mach. Learn. Appl. (ICCCMLA)*, Oct. 2022, pp. 534–539, doi: 10.1109/ICCCMLA56841.2022.9989289.
[87] P. Isaranontakul and W. Kreesuradej, ''A study of using GPT-3 to generate a Thai sentiment analysis of COVID-19 tweets dataset,'' in *Proc. 20th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jun. 2023, pp. 106–111, doi: 10.1109/jcsse58229.2023.10201994.
[88] N. Aydin and O. A. Erdem, ''A research on the new generation artificial intelligence technology generative pretraining transformer 3,'' in *Proc. 3rd Int. Informat. Softw. Eng. Conf. (IISEC)*, Dec. 2022, pp. 1–6, doi: 10.1109/IISEC56263.2022.9998298.
[89] M. Lajkó, V. Csuvik, and L. Vidács, ''Towards Javascript program repair with generative pre-trained transformer (GPT-2),'' in *Proc. IEEE/ACM Int. Workshop Automated Program Repair (APR)*, May 2022, pp. 61–68, doi: 10.1145/3524459.3527350.
[90] Y. Su, ''Computer-generated humour based on GPT-2,'' in *Proc. IEEE 2nd Int. Conf. Data Sci. Comput. Appl. (ICDSCA)*, Oct. 2022, pp. 890–893, doi: 10.1109/ICDSCA56264.2022.9987901.
[91] N. Darapaneni, R. Prajeesh, P. Dutta, V. K. Pillai, A. Karak, and A. R. Paduri, ''Abstractive text summarization using BERT and GPT-2 models,'' in *Proc. Int. Conf. Signal Process., Comput., Electron., Power Telecommun. (IConSCEPT)*, May 2023, pp. 1–6, doi: 10.1109/IConSCEPT57958.2023.10170093.
[92] Y. Liang and Z. Han, ''Intelligent love letter generator based on GPT-2 model,'' in *Proc. 3rd Int. Conf. Electron. Commun. Artif. Intell. (IWECAI)*, Jan. 2022, pp. 562–567, doi: 10.1109/IWECAI55315.2022.00115.
[93] D. H. Nguyen, H. L. Pham, and L. Le Thi Trang, ''Security of the cryptosystem GPT based on rank codes and term-rank codes,'' in *Proc. Int. Conf. Eng. Telecommun. (En&T)*, Nov. 2021, pp. 1–5, doi: 10.1109/EnT50460.2021.9681778.
[94] M. Nam, S. Park, and D. S. Kim, ''Intrusion detection method using bidirectional GPT for in-vehicle controller area networks,'' *IEEE Access*, vol. 9, pp. 124931–124944, 2021, doi: 10.1109/ACCESS.2021.3110524.
[95] D. Demirci, N. sahin, M. sirlancis, and C. Acarturk, ''Static malware detection using stacked BiLSTM and GPT-2,'' *IEEE Access*, vol. 10, pp. 58488–58502, 2022, doi: 10.1109/ACCESS.2022.3179384.
[96] H. Khan, M. Alam, S. Al-Kuwari, and Y. Faheem, ''Offensive AI: Unification of email generation through GPT-2 model with a game-theoretic approach for spear-phishing attacks,'' in *Proc. Competitive Advantage Digit. Economy (CADE)*, Jun. 2021, pp. 178–184, doi: 10.1049/ICP.2021.2422.
[97] J. J. Bird, M. Pritchard, A. Fratini, A. Ekárt, and D. R. Faria, ''Synthetic biological signals machine-generated by GPT-2 improve the classification of EEG and EMG through data augmentation,'' *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 3498–3504, Apr. 2021, doi: 10.1109/LRA.2021.3056355.
[98] P. Maddigan and T. Susnjak, ''Chat2VIS: Generating data visualizations via natural language using ChatGPT, Codex and GPT-3 large language models,'' *IEEE Access*, vol. 11, pp. 45181–45193, 2023, doi: 10.1109/ACCESS.2023.3274199.
[99] H. Liu, Y. Cai, Z. Lin, Z. Ou, Y. Huang, and J. Feng, ''Variational latent-state GPT for semi-supervised task-oriented dialog systems,'' *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 31, pp. 970–984, 2023, doi: 10.1109/TASLP.2023.3240661.
[100] S. W. Jeong, C. G. Kim, and T. K. Whangbo, ''Question answering system for healthcare information based on BERT and GPT,'' in *Proc. Joint Int. Conf. Digit. Arts, Media Technol. ECTI Northern Sect. Conf. Electr., Electron., Comput. Telecommun. Eng. (ECTI DAMT NCON)*, Mar. 2023, pp. 348–352, doi: 10.1109/ectidamtncon57770.2023.10139365.
[101] Y. Zhang, Z. Li, and J. Zhang, ''Towards the use of pretrained language model GPT-2 for testing the hypothesis of communicative efficiency in the lexicon,'' in *Proc. Int. Conf. Asian Lang. Process. (IALP)*, Dec. 2021, pp. 62–66, doi: 10.1109/IALP54817.2021.9675217.
[102] R. Kinoshita and S. Shiramatsu, ''Agent for recommending information relevant to web-based discussion by generating query terms using GPT-3,'' in *Proc. IEEE Int. Conf. Agents (ICA)*, Nov. 2022, pp. 24–29, doi: 10.1109/ICA55837.2022.00011.

[103] C. Treude, "Navigating complexity in software engineering: A prototype for comparing GPT-n solutions," in *Proc. IEEE/ACM 5th Int. Workshop Bots Softw. Eng. (BotSE)*, May 2023, pp. 1–5, doi: 10.1109/botse59190.2023.00008.

[104] J. Achiam et al., "GPT-4 technical report," 2023, *arXiv:2303.08774*.

[105] H. Nori, N. King, S. M. McKinney, D. Carignan, and E. Horvitz, "Capabilities of GPT-4 on medical challenge problems," 2023, *arXiv:2303.13375*.

[106] D. M. Katz, M. J. Bommarito, S. Gao, and P. Arredondo, "GPT-4 passes the bar exam," *SSRN Electron. J.*, Mar. 2023, doi: 10.2139/ssrn.4389233.

[107] D. Hendrycks, C. Burns, S. Basart, A. Zou, M. Mazeika, D. Song, and J. Steinhardt, "Measuring massive multitask language understanding," 2020, *arXiv:2009.03300*.

[108] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, Jan. 2018, doi: 10.1109/MSP.2017.2765202.

[109] A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 1, Apr. 2021, Art. no. 100004, doi: 10.1016/j.jjimei.2020.100004.

[110] Z. Zhang, M. Li, and J. Yu, "On the convergence and mode collapse of GAN," in *Proc. SIGGRAPH Asia Tech. Briefs*. New York, NY, USA: Association for Computing Machinery, Dec. 2018, pp. 1–4, doi: 10.1145/3283254.3283282.

[111] Bhagyashree, V. Kushwaha, and G. C. Nandi, "Study of prevention of mode collapse in generative adversarial network (GAN)," in *Proc. IEEE 4th Conf. Inf. Commun. Technol. (CICT)*, Dec. 2020, pp. 1–6, doi: 10.1109/CICT51604.2020.9312049.

[112] H. Thanh-Tung and T. Tran, "Catastrophic forgetting and mode collapse in GANs," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–10, doi: 10.1109/IJCNN48605.2020.9207181.

[113] W. Li, L. Fan, Z. Wang, C. Ma, and X. Cui, "Tackling mode collapse in multi-generator GANs with orthogonal vectors," *Pattern Recognit.*, vol. 110, Feb. 2021, Art. no. 107646, doi: 10.1016/j.patcog.2020.107646.

[114] D. Saxena and J. Cao, "Generative adversarial networks (GANs): Challenges, solutions, and future directions," *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1–42, 2021.

[115] H. Chen, "Challenges and corresponding solutions of generative adversarial networks (GANs): A survey study," *J. Phys., Conf. Ser.*, vol. 1827, no. 1, Mar. 2021, Art. no. 012066, doi: 10.1088/1742-6596/1827/1/012066.

[116] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[117] G. G. Chrysos, J. Kossaifi, and S. Zafeiriou, "Robust conditional generative adversarial networks," 2019, *arXiv:1805.08657*.

[118] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.

[119] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*.

[120] T. C. Koopmans, "Optimum utilization of the transportation system," *Econometrica*, vol. 17, p. 136, Jul. 1949, doi: 10.2307/1907301.

[121] E. Massart, "Improving weight clipping in Wasserstein GANs," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*. Montreal, QC, Canada: IEEE, Aug. 2022, pp. 2286–2292, doi: 10.1109/ICPR56361.2022.9956056.

[122] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251. Accessed: Aug. 27, 2023. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html

[123] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8789–8797. Accessed: Aug. 27, 2023.

[124] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, *arXiv:1710.10196*.

[125] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," 2018, *arXiv:1809.11096*.

[126] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4396–4405. Accessed: Aug. 29, 2023. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/html/Karras_A_Style-Based_Generator_Architecture_for_Generative_Adversarial_Networks_CVPR_2019_paper.html

[127] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 8107–8116. Accessed: Aug. 29, 2023. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2020/html/Karras_Analyzing_and_Improving_the_Image_Quality_of_StyleGAN_CVPR_2020_paper.html

[128] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel, "InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2016. Accessed: Aug. 29, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2016/hash/7c9d0b1f96aebd7b5eca8c3edaa19ebb-Abstract.html

[129] T. Salimans et al., "Improved techniques for training GANs," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2016. Accessed: Aug. 29, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2016/hash/8a3363abe792db2d8761d6403605aeb7-Abstract.html

[130] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," 2017, *arXiv:1605.09782*.

[131] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion models: A comprehensive survey of methods and applications," *ACM Comput. Surv.*, vol. 56, no. 4, pp. 1–39, Nov. 2023, doi: 10.1145/3626235.

[132] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li, "A survey on generative diffusion models," *IEEE Trans. Knowl. Data Eng.*, 2024, doi: 10.1109/TKDE.2024.3361474.

[133] Y. Pang, J. Mao, L. He, H. Lin, and Z. Qiang, "An improved face image restoration method based on denoising diffusion probabilistic models," *IEEE Access*, vol. 12, pp. 3581–3596, 2024, doi: 10.1109/ACCESS.2024.3349423.

[134] F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10850–10869, Sep. 2023, doi: 10.1109/TPAMI.2023.3261988.

[135] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2020, pp. 6840–6851. Accessed: Sep. 4, 2023. [Online]. Available: https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html

[136] Q. Yi, X. Chen, C. Zhang, Z. Zhou, L. Zhu, and X. Kong, "Diffusion models in text generation: A survey," *PeerJ Comput. Sci.*, vol. 10, p. e1905, Feb. 2024, doi: 10.7717/peerj-cs.1905.

[137] Y. Song and S. Ermon, "Improved techniques for training score-based generative models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 12438–12448.

[138] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2019. Accessed: Sep. 4, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2019/hash/3001ef257407d5a371a96dcd947c7d93-Abstract.html

[139] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2020, *arXiv:2011.13456*.

[140] B. D. O. Anderson, "Reverse-time diffusion equation models," *Stochastic Processes Appl.*, vol. 12, no. 3, pp. 313–326, May 1982, doi: 10.1016/0304-4149(82)90051-5.

[141] Deloitte. (Jan. 2024). *U.S. State of Gen AI Report*. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-state-of-gen-ai-report.pdf

[142] P. Denny, J. Leinonen, J. Prather, A. Luxton-Reilly, T. Amarouche, B. A. Becker, and B. N. Reeves, "Prompt problems: A new programming exercise for the generative AI era," in *Proc. 55th ACM Tech. Symp. Comput. Sci. Educ.* Portland, OR, USA: ACM, Mar. 2024, pp. 296–302, doi: 10.1145/3626252.3630909.

[143] *Prompt Engineering Guide*. Accessed: Mar. 29, 2024. [Online]. Available: https://www.promptingguide.ai/techniques/zeroshot
[144] *OpenAI Platform*. Accessed: Mar. 29, 2024. [Online]. Available: https://platform.openai.com
[145] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, and M. Chen, "Hierarchical text-conditional image generation with CLIP latents," 2022, *arXiv:2204.06125*.
[146] J. Ho et al., "Imagen video: High definition video generation with diffusion models," 2022, *arXiv:2210.02303*.
[147] U. Singer et al., "Make-A-video: Text-to-video generation without text-video data," 2022, *arXiv:2209.14792*.
[148] Y. Li et al., "Competition-level code generation with AlphaCode," *Science*, vol. 378, no. 6624, pp. 1092–1097, Dec. 2022, doi: 10.1126/science.abq1158.
[149] *StarCoder: A State-of-the-Art LLM for Code*. Accessed: Oct. 6, 2023. [Online]. Available: https://huggingface.co/blog/starcoder
[150] M. Chen et al., "Evaluating large language models trained on code," 2021, *arXiv:2107.03374*.
[151] M. Civit, J. Civit-Masot, F. Cuadrado, and M. J. Escalona, "A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends," *Expert Syst. Appl.*, vol. 209, Dec. 2022, Art. no. 118190. https://www.sciencedirect.com/science/article/pii/S0957417422013537
[152] *MuseNet*. Accessed: Oct. 7, 2023. [Online]. Available: https://openai.com/research/musenet
[153] P. Dhariwal, H. Jun, C. Payne, J. W. Kim, A. Radford, and I. Sutskever, "Jukebox: A generative model for music," 2020, *arXiv:2005.00341*.
[154] House Beautiful. *How to Give Your Living Room a Luxurious Makeover*. Accessed: Sep. 26, 2023. [Online]. Available: https://www.housebeautiful.com/room-decorating/living-family-rooms/g715/designer-living-rooms/
[155] Y. Yoon, B. Cha, J.-H. Lee, M. Jang, J. Lee, J. Kim, and G. Lee, "Speech gesture generation from the trimodal context of text, audio, and speaker identity," *ACM Trans. Graph.*, vol. 39, no. 6, pp. 1–16, Nov. 2020, doi: 10.1145/3414685.3417838.
[156] Y. Ning, S. He, Z. Wu, C. Xing, and L.-J. Zhang, "A review of deep learning based speech synthesis," *Appl. Sci.*, vol. 9, no. 19, p. 4050, Sep. 2019, doi: 10.3390/app9194050.
[157] M. Le et al., "Voicebox: Text-guided multilingual universal speech generation at scale," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 36, Dec. 2023, pp. 14005–14034.
[158] N. Kaur and P. Singh, "Conventional and contemporary approaches used in text to speech synthesis: A review," *Artif. Intell. Rev.*, vol. 56, no. 7, pp. 5837–5880, Jul. 2023, doi: 10.1007/s10462-022-10315-0.
[159] A. Wali, Z. Alamgir, S. Karim, A. Fawaz, M. B. Ali, M. Adan, and M. Mujtaba, "Generative adversarial networks for speech processing: A review," *Comput. Speech Lang.*, vol. 72, Mar. 2022, Art. no. 101308, doi: 10.1016/j.csl.2021.101308.
[160] J. A Rodriguez, D. Vazquez, I. Laradji, M. Pedersoli, and P. Rodriguez, "FigGen: Text to scientific figure generation," 2023, *arXiv:2306.00800*.
[161] *Minerva: Solving Quantitative Reasoning Problems With Language Models*. Accessed: Oct. 7, 2023. [Online]. Available: https://blog.research.google/2022/06/minerva-solving-quantitative-reasoning.html
[162] R. Taylor, M. Kardas, G. Cucurull, T. Scialom, A. Hartshorn, E. Saravia, A. Poulton, V. Kerkez, and R. Stojnic, "Galactica: A large language model for science," 2022, *arXiv:2211.09085*.
[163] *Research Blog: DeepDream—A Code Example for Visualizing Neural Networks*. Accessed: Apr. 6, 2024. [Online]. Available: http://googleresearch.blogspot.co.uk/2015/07/deepdream-code-example-for-visualizing.html
[164] Google DeepMind. *WaveNet: A Generative Model for Raw Audio*. Accessed: Apr. 6, 2024. [Online]. Available: https://deepmind.google/discover/blog/wavenet-a-generative-model-for-raw-audio/
[165] (2016). *Adobe Sensei Lets Customers Master the Art of Digital Experiences*. Accessed: Apr. 6, 2024. [Online]. Available: https://news.adobe.com/news/news-details/2016/Adobe-Sensei-Lets-Customers-Master-the-Art-of-Digital-Experiences/default.aspx
[166] (2024). *Best AI Video Generator in 2024—Synthesia*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.synthesia.io/
[167] A. J. Mitchell. *The All-New Boomy: Make Instant Music, Reach Billions*. Accessed: Apr. 6, 2024. [Online]. Available: https://blog.boomy.com/the-all-new-boomy-make-instant-music-reach-billions-3bab562a75c5
[168] *AI Art Generator: Create Stunning AI Art*. Accessed: Apr. 6, 2024. [Online]. Available: https://creator.nightcafe.studio/
[169] *Descript | All-in-One Video & Podcast Editing, Easy as a Doc*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.descript.com/
[170] *AI Voice Generator: Realistic Text to Speech & Voice Cloning*. Accessed: Apr. 6, 2024. [Online]. Available: https://lovo.ai/
[171] *AI Voice Generator: Versatile Text to Speech Software | Murf AI*. Accessed: Apr. 6, 2024. [Online]. Available: https://murf.ai
[172] *Unlock Generative AI Content At Scale*. Accessed: Apr. 6, 2024. [Online]. Available: https://synthesys.io/
[173] *AI Art Generator From Text*. Accessed: Apr. 6, 2024. [Online]. Available: https://writesonic.com/photosonic-ai-art-generator
[174] *NovelAI—The AI Storyteller*. Accessed: Apr. 6, 2024. [Online]. Available: https://novelai.net
[175] *Overdub*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.descript.com/overdub-fb
[176] *Jasper | AI Copilot for Enterprise Marketing Teams*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.jasper.ai/
[177] N. Friedman. (2021). *Introducing GitHub Copilot: Your AI Pair Programmer*. Accessed: Apr. 6, 2024. [Online]. Available: https://github.blog/2021-06-29-introducing-github-copilot-ai-pair-programmer/
[178] *OpenAI Codex*. Accessed: Apr. 6, 2024. [Online]. Available: https://openai.com/blog/openai-codex
[179] *Automated Meeting Reports, Transcripts, Notes, Video, Coaching, Scheduling | Read*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.read.ai/
[180] *#1 AI Music Studio—AI Music Generator for Creators*. Accessed: Apr. 6, 2024. [Online]. Available: https://soundful.com/en-us/
[181] *SOUNDRAW*. Accessed: Apr. 6, 2024. [Online]. Available: https://soundraw.io
[182] *AI Art Generator App—AI Art Maker*. Accessed: Apr. 6, 2024. [Online]. Available: https://starryai.com/
[183] *Textalky—AI Text to Speech Synthesis Software*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.textalky.com/
[184] *Generate*. Accessed: Apr. 6, 2024. [Online]. Available: https://cohere.com/generate
[185] L. Cai, Y. Chen, N. Cai, W. Cheng, and H. Wang, "Utilizing Amari-alpha divergence to stabilize the training of generative adversarial networks," *Entropy*, vol. 22, no. 4, p. 410, Apr. 2020, doi: 10.3390/e22040410.
[186] Google. *LaMDA: Our Breakthrough Conversation Technology*. Accessed: Apr. 6, 2024. [Online]. Available: https://blog.google/technology/ai/lamda/
[187] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E. Denton, S. K. S. Ghasemipour, B. K. Ayan, S. S. Mahdavi, R. G. Lopes, T. Salimans, J. Ho, D. J Fleet, and M. Norouzi, "Photorealistic text-to-image diffusion models with deep language understanding," 2022, *arXiv:2205.11487*.
[188] T. Schick et al., "PEER: A collaborative language model," 2022, *arXiv:2208.11663*.
[189] *Midjourney Documentation and User Guide*. Accessed: Apr. 6, 2024. [Online]. Available: https://docs.midjourney.com
[190] *Perplexity*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.perplexity.ai/
[191] *Introducing Whisper*. Accessed: Apr. 6, 2024. [Online]. Available: https://openai.com/research/whisper
[192] *Stable Diffusion Public Release*. Accessed: Apr. 6, 2024. [Online]. Available: https://stability.ai/news/stable-diffusion-public-release
[193] *Stability AI Releases StableStudio, the Open Source Future of DreamStudio*. Accessed: Apr. 6, 2024. [Online]. Available: https://stability.ai/news/stablestudio-open-source-community-driven-future-dreamstudio-release
[194] *Chatsonic—Free AI Chat | GPT-4 & ChatGPT Alternative*. Accessed: Apr. 6, 2024. [Online]. Available: https://writesonic.com/chat
[195] (2024). *Adobe Firefly Feature Summary*. Accessed: Apr. 6, 2024. [Online]. Available: https://helpx.adobe.com/content/help/en/firefly/using/whats-new/2024.html
[196] *Wordtune | Free Generative AI Writing Tool*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.wordtune.com/lp/aigenerative?utm_source=google-search&utm_medium=cpc&utm_campaign_id=20308562145&ad_set_id=149915183479&utm_campaign=Write_Search_US_EN_D_SUPS_tCPA_Brand_Main&ad_set_name=Wordtune_Search_US_EN_D_Brand_Main&utm_term=word

[197] *Blog*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.altered.ai/blog/
[198] *Amper | Music for Your Defining Moment | Welcome AI*. Accessed: Apr. 6, 2024. [Online]. Available: https://welcome.ai/solution/amper
[199] *Claude 2*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.anthropic.com/news/claude-2
[200] (2023). *Apple Preps Ajax Generative AI, 'Apple GPT' to Rival OpenAI and Google—Bloomberg*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.bloomberg.com/news/articles/2023-07-19/apple-preps-ajax-generative-ai-apple-gpt-to-rival-openai-and-google
[201] Inflection. *Press Release: Inflection AI Introduces Pi, Your Personal AI*. Accessed: Apr. 6, 2024. [Online]. Available: https://inflection.ai/press
[202] Canva. *Free AI Image Generator: Online Text to Image App*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.canva.com/ai-image-generator/
[203] *Introducing PaLM 2*. Accessed: Apr. 6, 2024. [Online]. Available: https://blog.google/technology/ai/google-palm-2-ai-large-language-model/
[204] Google Cloud Blog. *Google Cloud Launches New AI Models, Opens Generative AI Studio*. Accessed: Apr. 6, 2024. [Online]. Available: https://cloud.google.com/blog/products/ai-machine-learning/google-cloud-launches-new-ai-models-opens-generative-ai-studio
[205] *An Important Next Step on Our AI Journey*. Accessed: Apr. 6, 2024. [Online]. Available: https://blog.google/technology/ai/bard-google-ai-search-updates/
[206] R. Villegas, M. Babaeizadeh, P.-J. Kindermans, H. Moraldo, H. Zhang, M. T. Saffar, S. Castro, J. Kunze, and D. Erhan, "Phenaki: Variable length video generation from open domain textual description," 2022, *arXiv:2210.02399*.
[207] J. Yu, Y. Xu, J. Yu Koh, T. Luong, G. Baid, Z. Wang, V. Vasudevan, A. Ku, Y. Yang, B. K. Ayan, B. Hutchinson, W. Han, Z. Parekh, X. Li, H. Zhang, J. Baldridge, and Y. Wu, "Scaling autoregressive models for content-rich text-to-image generation," 2022, *arXiv:2206.10789*.
[208] *Turn Ideas Into Music With MusicLM*. Accessed: Apr. 6, 2024. [Online]. Available: https://blog.google/technology/ai/musiclm-google-ai-test-kitchen/
[209] Android Developers. *Meet Studio Bot | Android Studio*. Accessed: Apr. 6, 2024. [Online]. Available: https://developer.android.com/studio/preview/studio-bot
[210] *VoiceBox*. Accessed: Apr. 6, 2024. [Online]. Available: https://voicebox.dev.metademolab.com/
[211] *New AI Experiences and Tools | Meta AI*. Accessed: Apr. 6, 2024. [Online]. Available: https://ai.meta.com/genai
[212] T. Wang, P. Yu, X. Ellen Tan, S. O'Brien, R. Pasunuru, J. Dwivedi-Yu, O. Golovneva, L. Zettlemoyer, M. Fazel-Zarandi, and A. Celikyilmaz, "Shepherd: A critic for language model generation," 2023, *arXiv:2308.04592*.
[213] *Introducing CM3leon, a More Efficient, State-of-the-Art Generative Model for Text and Images*. Accessed: Apr. 6, 2024. [Online]. Available: https://ai.meta.com/blog/generative-ai-text-images-cm3leon/
[214] *AudioCraft*. Accessed: Apr. 6, 2024. [Online]. Available: https://ai.meta.com/resources/models-and-libraries/audiocraft
[215] *MusicGen—A Hugging Face Space by Facebook*. Accessed: Apr. 6, 2024. [Online]. Available: https://huggingface.co/spaces/facebook/MusicGen
[216] J. Betker et al., "Improving image generation with better captions," *OpenAI*, vol. 2, no. 3, p. 8, 2023.
[217] *RoomGPT—Redesign Your Room With AI*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.roomgpt.io/
[218] *Gen-2*. Accessed: Apr. 6, 2024. [Online]. Available: https://runwayml.com/ai-tools/gen-2/
[219] (2023). *Salesforce Announces Einstein GPT, the World's First Generative AI for CRM—Salesforce*. Accessed: Apr. 6, 2024. [Online]. Available: https://www.salesforce.com/news/press-releases/2023/03/07/einstein-generative-ai/
[220] *Scribe—Powered by AI*. Accessed: Apr. 6, 2024. [Online]. Available: https://scribehow.com/scribe-ai
[221] *Introducing Devin, the First AI Software Engineer*. Accessed: Apr. 7, 2024. [Online]. Available: https://www.cognition-labs.com/introducing-devin
[222] *Sora: Creating Video From Text*. Accessed: Apr. 7, 2024. [Online]. Available: https://openai.com/sora
[223] *Text to Speech Software, Transforming Text to Voice Instantly*. Accessed: Apr. 7, 2024. [Online]. Available: https://speechelo.com/
[224] *Kits AI—Studio-quality AI Music Tools*. Accessed: Apr. 7, 2024. [Online]. Available: https://www.kits.ai/
[225] *The #1 AI Voice Over Generator*. Accessed: Apr. 7, 2024. [Online]. Available: https://speechify.com/voiceover-lp-3/?utm_source=googlevoiceover&utm_medium=cpc&utm_campaign=voiceover&utm_content=voiceover&utm_term=speechify
[226] *Paraphrasing Tool—QuillBot AI*. Accessed: Apr. 7, 2024. [Online]. Available: https://quillbot.com
[227] WellSaid Labs. *Home*. Accessed: Apr. 7, 2024. [Online]. Available: https://wellsaidlabs.com/
[228] *Generative AI Workspace for Visual Productivity*. Accessed: Apr. 7, 2024. [Online]. Available: https://www.jeda.ai/
[229] C. Plut and P. Pasquier, "Generative music in video games: State of the art, challenges, and prospects," *Entertainment Comput.*, vol. 33, Mar. 2020, Art. no. 100337, doi: 10.1016/j.entcom.2019.100337.
[230] S. Wang, S. Menon, T. Long, K. Henderson, D. Li, K. Crowston, M. Hansen, J. V. Nickerson, and L. B. Chilton, "ReelFramer: Human-AI co-creation for news-to-video translation," 2023, *arXiv:2304.09653*.
[231] T. H. D. and R. Bean. *The Impact of Generative AI on Hollywood and Entertainment*. MIT Sloan Manage. Rev. Accessed: Oct. 11, 2023. [Online]. Available: https://sloanreview.mit.edu/article/the-impact-of-generative-ai-on-hollywood-and-entertainment/
[232] *Runway AI: Tech Behind Everything Everywhere All at Once*. Accessed: Oct. 11, 2023. [Online]. Available: https://topten.ai/ai-tech-behind-everything-everywhere-all-at-once/
[233] L. Rai, C. Deng, and F. Liu, "Developing massive open online course style assessments using generative AI tools," in *Proc. IEEE 6th Int. Conf. Electron. Inf. Commun. Technol. (ICEICT)*, Jul. 2023, pp. 1292–1294, doi: 10.1109/iceict57916.2023.10244824.
[234] W. M. Lim, A. Gunasekara, J. L. Pallant, J. I. Pallant, and E. Pechenkina, "Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators," *Int. J. Manage. Educ.*, vol. 21, no. 2, Jul. 2023, Art. no. 100790, doi: 10.1016/j.ijme.2023.100790.
[235] Y.-X. Li and N.-C. Tai, "Teaching at the right moment: A generative AI-enabled bedtime storybook generation system communicating timely issues," in *Proc. Int. Conf. Consum. Electron.-Taiwan (ICCE-Taiwan)*, Jul. 2023, pp. 157–158, doi: 10.1109/icce-taiwan58799.2023.10226626.
[236] E. A. Alasadi and C. R. Baiz, "Generative AI in education and research: Opportunities, concerns, and solutions," *J. Chem. Educ.*, vol. 100, no. 8, pp. 2965–2971, Aug. 2023, doi: 10.1021/acs.jchemed.3c00323.
[237] H. Yu, Z. Liu, and Y. Guo, "Application status, problems and future prospects of generative AI in education," in *Proc. 5th Int. Conf. Comput. Sci. Technol. Educ. (CSTE)*, Apr. 2023, pp. 1–7, doi: 10.1109/cste59648.2023.00065.
[238] M. Kuzlu, Z. Xiao, S. Sarp, F. O. Catak, N. Gurler, and O. Guler, "The rise of generative artificial intelligence in healthcare," in *Proc. 12th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2023, pp. 1–4, doi: 10.1109/MECO58584.2023.10155107.
[239] A. Jadon and S. Kumar, "Leveraging generative AI models for synthetic data generation in healthcare: Balancing research and privacy," in *Proc. Int. Conf. Smart Appl., Commun. Netw. (SmartNets)*, Jul. 2023, pp. 1–4, doi: 10.1109/smartnets58706.2023.10215825.
[240] Z. Shen, F. Ding, A. Jolfaei, K. Yadav, S. Vashisht, and K. Yu, "DeformableGAN: Generating medical images with improved integrity for healthcare cyber physical systems," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 5, pp. 2584–2596, Sep. 2023, doi: 10.1109/TNSE.2022.3190765.
[241] (2023). *AWS Announces AWS HealthScribe, a New Generative AI-Powered Service That Automatically Creates Clinical Documentation*. Accessed: Oct. 11, 2023. [Online]. Available: https://press.aboutamazon.com/2023/7/aws-announces-aws-healthscribe-a-new-generative-ai-powered-service-that-automatically-creates-clinical-documentation
[242] B. Tang, J. Ewalt, and H.-L. Ng, "Generative AI models for drug discovery," in *Biophysical and Computational Tools in Drug Discovery* (Topics in Medicinal Chemistry), A. K. Saxena, Ed. Cham, Switzerland: Springer, 2021, pp. 221–243, doi: 10.1007/7355_2021_124.
[243] W. P. Walters and M. Murcko, "Assessing the impact of generative AI on medicinal chemistry," *Nature Biotechnol.*, vol. 38, no. 2, pp. 143–145, Feb. 2020, doi: 10.1038/s41587-020-0418-2.

[244] X. Zeng, F. Wang, Y. Luo, S.-G. Kang, J. Tang, F. C. Lightstone, E. F. Fang, W. Cornell, R. Nussinov, and F. Cheng, "Deep generative molecular design reshapes drug discovery," *Cell Rep. Med.*, vol. 3, no. 12, Dec. 2022, Art. no. 100794, doi: 10.1016/j.xcrm.2022.100794.

[245] A. Madani, B. Krause, E. R. Greene, S. Subramanian, B. P. Mohr, J. M. Holton, J. L. Olmos, C. Xiong, Z. Z. Sun, R. Socher, J. S. Fraser, and N. Naik, "Large language models generate functional protein sequences across diverse families," *Nature Biotechnol.*, vol. 41, no. 8, pp. 1099–1106, Aug. 2023, doi: 10.1038/s41587-022-01618-2.

[246] E. Sevgen et al., "ProT-VAE: Protein transformer variational AutoEncoder for functional protein design," *BioRxiv*, Jan. 2023, doi: 10.1101/2023.01.23.525232.

[247] (2023). *Cognizant and Google Cloud Expand Alliance to Bring AI to Enterprise Clients*. Accessed: Oct. 11, 2023. [Online]. Available: https://news.cognizant.com/2023-05-09-Cognizant-and-Google-Cloud-Expand-Alliance-to-Bring-AI-to-Enterprise-Clients

[248] (2032). *Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds | Press | Bloomberg LP*. Accessed: Oct. 12, 2023. [Online]. Available: https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

[249] T. H. Baek, "Digital advertising in the age of generative AI," *J. Current Issues Res. Advertising*, vol. 44, no. 3, pp. 249–251, Jul. 2023, doi: 10.1080/10641734.2023.2243496.

[250] J. Huh, M. R. Nelson, and C. A. Russell, "ChatGPT, AI advertising, and advertising research and education," *J. Advertising*, vol. 52, no. 4, pp. 477–482, Aug. 2023, doi: 10.1080/00913367.2023.2227013.

[251] J. Ford, V. Jain, K. Wadhwani, and D. G. Gupta, "AI advertising: An overview and guidelines," *J. Bus. Res.*, vol. 166, Nov. 2023, Art. no. 114124, doi: 10.1016/j.jbusres.2023.114124.

[252] A. Beheshti, J. Yang, Q. Z. Sheng, B. Benatallah, F. Casati, S. Dustdar, H. R. M. Nezhad, X. Zhang, and S. Xue, "ProcessGPT: Transforming business process management with generative artificial intelligence," in *Proc. IEEE Int. Conf. Web Services (ICWS)*, Jul. 2023, pp. 731–739, doi: 10.1109/icws60048.2023.00099.

[253] *Amazon Launches Generative AI to Help Sellers Write Product Descriptions*. Accessed: Oct. 12, 2023. [Online]. Available: https://www.aboutamazon.com/news/small-business/amazon-sellers-generative-ai-tool

[254] M. A. Ferrag, D. Hamouda, M. Debbah, L. Maglaras, and A. Lakas, "Generative adversarial networks-driven cyber threat intelligence detection framework for securing Internet of Things," in *Proc. 19th Int. Conf. Distrib. Comput. Smart Syst. Internet Things (DCOSS-IoT)*, Jun. 2023, pp. 196–200, doi: 10.1109/dcoss-iot58021.2023.00042.

[255] J.-H. Oh, D.-J. Lee, C.-H. Ji, D.-H. Shin, J.-W. Han, Y.-H. Son, and T.-E. Kam, "Graph-based conditional generative adversarial networks for major depressive disorder diagnosis with synthetic functional brain network generation," *IEEE J. Biomed. Health Informat.*, vol. 28, no. 3, pp. 1504–1515, Mar. 2024, doi: 10.1109/JBHI.2023.3340325.

[256] R. K. Yadav, S. Gupta, and P. D, "Real-time anomaly detection with generative adversarial networks for network security management," in *Proc. Int. Conf. Optim. Comput. Wireless Commun. (ICOCWC)*, Jan. 2024, pp. 1–6, doi: 10.1109/icocwc60930.2024.10470902.

[257] H. Ding, Y. Sun, N. Huang, Z. Shen, and X. Cui, "TMG-GAN: Generative adversarial networks-based imbalanced learning for network intrusion detection," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 1156–1167, 2024, doi: 10.1109/TIFS.2023.3331240.

[258] C. Park, J. Lee, Y. Kim, J.-G. Park, H. Kim, and D. Hong, "An enhanced AI-based network intrusion detection system using generative adversarial networks," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2330–2345, Feb. 2023, doi: 10.1109/JIOT.2022.3211346.

[259] A. Celik and A. M. Eltawil, "At the dawn of generative AI era: A tutorial-cum-survey on new frontiers in 6G wireless intelligence," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2433–2489, 2024, doi: 10.1109/OJCOMS.2024.3362271.

[260] *Jobs of Tomorrow: Large Language Models and Jobs*. Accessed: Oct. 14, 2023. [Online]. Available: https://www.weforum.org/whitepapers/jobs-of-tomorrow-large-language-models-and-jobs/

[261] *Automation or Augmentation? This is How AI Will be Integrated Into the Jobs of Tomorrow*. Accessed: Oct. 14, 2023. [Online]. Available: https://www.weforum.org/agenda/2023/09/ai-automation-augmentation-workplace-jobs-of-tomorrow/

[262] *We Always Hear That AI Will Take Our Jobs. But What Jobs Will it Create?*. Accessed: Oct. 14, 2023. [Online]. Available: https://www.weforum.org/agenda/2023/09/jobs-ai-will-create/

[263] K. Michael, R. Abbas, and G. Roussos, "AI in cybersecurity: The paradox," *IEEE Trans. Technol. Soc.*, vol. 4, no. 2, pp. 104–109, Jun. 2023, doi: 10.1109/TTS.2023.3280109.

[264] S. Oh and T. Shon, "Cybersecurity issues in generative AI," in *Proc. Int. Conf. Platform Technol. Service (PlatCon)*, Aug. 2023, pp. 97–100, doi: 10.1109/platcon60102.2023.10255179.

[265] M. Gupta, C. Akiri, K. Aryal, E. Parker, and L. Praharaj, "From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy," *IEEE Access*, vol. 11, pp. 80218–80245, 2023, doi: 10.1109/ACCESS.2023.3300381.

[266] *AI-Based Cybercrime Tools WormGPT and FraudGPT Could be The Tip of the Iceberg | SlashNext*. Accessed: Aug. 24, 2023. [Online]. Available: https://slashnext.com/blog/ai-based-cybercrime-tools-wormgpt-and-fraudgpt-could-be-the-tip-of-the-iceberg/

[267] I.-C. Mihai, "The transformative impact of artificial intelligence on cybersecurity," *Int. J. Inf. Secur. Cybercrime*, vol. 12, no. 1, p. 9, 2023.

[268] P. V. Falade, "Decoding the threat landscape: ChatGPT, FraudGPT, and WormGPT in social engineering attacks," *Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol.*, pp. 185–198, Oct. 2023, doi: 10.32628/cseit2390533.

[269] M. Mozes, X. He, B. Kleinberg, and L. D. Griffin, "Use of LLMs for illicit purposes: Threats, prevention measures, and vulnerabilities," 2023, *arXiv:2308.12833*.

[270] *Pause Giant AI Experiments: An Open Letter*. Accessed: Aug. 24, 2023. [Online]. Available: https://futureoflife.org/open-letter/pause-giant-ai-experiments/

[271] J. Coscarelli. *An A.I. Hit of Fake 'Drake' and 'The Weeknd' Rattles the Music World*. The New York Times. Accessed: Oct. 11, 2023. [Online]. Available: https://www.nytimes.com/2023/04/19/arts/music/ai-drake-the-weeknd-fake.html

[272] B. Lane. (2023). *An AI-Generated Rihanna Cover of Beyonc?'s 'Cuff It' is Going Viral, and It Could Open Up a New Legal Nightmare for the Music Industry*. Insider. Accessed: Oct. 14, 2023. [Online]. Available: https://www.insider.com/rihanna-ai-cuff-it-cover-legal-nightmare-music-industry-2023-4

[273] Med. Futurist. *ChatGPT In Healthcare: What Science Says*. Accessed: Aug. 24, 2023. [Online]. Available: https://medicalfuturist.com/chatgpt-in-healthcare-what-the-science-says/

[274] *Generative AI and the Future of Work in America | McKinsey*. Accessed: Oct. 10, 2023. [Online]. Available: https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america

[275] A. Zarifhonarvar, "Economics of ChatGPT: A labor market view on the occupational impact of artificial intelligence," *SSRN J.*, Feb. 7, 2023, doi: 10.2139/ssrn.4350925.

[276] *Future of Work Report: AI at Work*. Accessed: Oct. 14, 2023. [Online]. Available: https://economicgraph.linkedin.com/research/future-of-work-report-ai

[277] Q. Hamirani. (2024). *Here's How Generative AI Will Redefine The Workplace*. Forbes. Accessed: Apr. 6, 2024. [Online]. Available: https://www.forbes.com/sites/qhamirani/2024/02/15/heres-how-generative-ai-will-redefine-the-workplace/

STAPHORD BENGESI (Graduate Student Member, IEEE) received the B.S. degree in computer science from the University of Dar es Salaam (UDSM), Dar es Salaam, Tanzania, in 2014, and the M.S. degree in computer science from Bowie State University, MD, USA, in 2021, where he is currently pursuing the Ph.D. degree in computer science.

From 2019 to 2021, he was a Research Graduate Assistant with the Computer Science Department, Bowie State University, where he has been an Adjunct Professor with the Computer Science Department, since 2021. His research interests include machine learning, deep learning, and cloud computing.

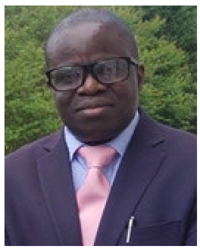
HODA EL-SAYED received the D.Sc. degree in computer science from The George Washington University.

She is currently a Professor with the Department of Computer Science, Bowie State University. Her research interests include parallel and distributed algorithms, high-performance computing, and programming. Prior to joining Bowie State University, she was a Researcher with the High-Performance Computing Division, National Institute of Standards and Technology (NIST). She has published many refereed research articles and her publications appeared in prestigious venues, such the IEEE Transactions and the IEEE International Parallel and Distributed Processing Symposium (IPDPS). Her research has been supported by Federal Agencies, including the National Security Agency and Department of Defense. She is also the Founding Director of the Center for High-Performance Information Processing (CHIP), a BSU interdisciplinary high-performance computing research center, which is the home to the BSU state-of-the-art Cray-XC-40 Supercomputer (Sphinx).

MD KAMRUZZAMAN (ZAMAN) SARKER is an Assistant Professor at Bowie State University. He has previous experience as an Assistant Professor at the University of Hartford. Before joining the University of Hartford, he was a Post-Doctoral Fellow at Kansas State University. He also worked for several industries, including Intel, Accenture, and Samsung. He love to invent new stuff (research), share knowledge (teaching), and at the same time build stuff (software engineering). His research domain spans the broad field of artificial intelligence (AI) and its application. Some specific topics of focus include but are not limited to explainable artificial intelligence, deep learning, knowledge graph, semantic web, and cybersecurity.

YAO HOUKPATI received the M.A. degree in urban sociology from the University of Lome, Togo, in 2008, and the M.A. degree in computational social science from George Mason University, VA, USA, in 2018. He is currently pursuing the D.S. degree in computer science with Bowie State University, MD, USA. He is also an Adjunct Professor with Bowie State University. His research interests include machine learning, deep learning, cloud computing, cybersecurity, and agent-based modeling.

JOHN IRUNGU received the M.S. degree in computer science. He is currently pursuing the Ph.D. degree in computer science and engineering with the University of the District of Columbia, Washington, DC, USA. His professional experience includes the academia and public sector, where he was a Data Analyst with the Government of the District of Columbia. Prior to that role, he was a Data Science Graduate Research Assistant with the DC City Administrator's Office as a first recipient of the University of the District of Columbia Research Council Scholarship, from 2018 to 2019. His research interests include application of artificial intelligence in acoustics, signal processing, and information security.



TIMOTHY OLADUNNI received the master's and Ph.D. degrees in computer science from Bowie State University, MD, USA, in 2013 and 2017, respectively.

He is currently an Assistant Professor in computer science with Morgan State University, MD, USA. Before joining Morgan State University, he was an Assistant Professor with the Department of Computer Science and Information Technology, University of the District of Columbia, Washington, DC, USA. He explores computer science fundamental concepts in developing sustainable, efficient, and innovative solutions to real world problem. He has a broad research experience in artificial intelligence with specific expertise in natural language processing, computer vision, data science, pattern recognition, and computational epidemiology.