ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 2, 2024

Exploring the Landscape of Generative AI: A Comprehensive Survey of Use Cases

Mr.K.Appa Rao

Assistant Professor, Department of CSE,
Malla Reddy College of Engineering for Women.,

Maisammaguda., Medchal., TS, India (@gmail.com

Abstract: In recent years, generative AI has grown remarkably, opening up a wide range of applications in several fields. With a structured taxonomy and succinct explanations of several unimodal and even multimodal generative AIs, we give a thorough assessment of over 350 generative AI applications in this work. Text (ChatGPT or Bard), pictures (DALL-E 3 or Midjourney), video (Colossyan AI and Synthesia AI), games (CSM and ConvAI), and brain (Meta's speech from brain) information are just a few of the unimodal generative AI applications that are included in the survey's parts. To the best of our knowledge, no previous work has provided an updated perspective of the generative AGI landscape with many examples of each technology. Other surveys have not gone into great into about the technology or models to utilize, instead concentrating on examining how generative has affected certain businesses. For each area, examples were gathered through a comprehensive online and industry investigation. This study was developed with ethical considerations in mind.

Keywords: deep learning, machine learning, generative AI, artificial intelligence, and AI

Introduction

It is believed that generative AI technologies will mark a sea change in technological advancement. Historically, categorization activities like fraud detection, client segmentation, and risk assessment have been effectively handled by artificial intelligence. The ability to produce new output that differs from training data and fosters creativity is the breakthrough with generative AI (Feuerriegel et al., 2024). Unlike expert systems, which only analyze or act on material that already exists, generative AI is capable of creating original content (Murphy, 2022). With the help of large data sets and complex architectures, generative AI models are remarkably capable of producing original and varied content. They are able to process and learn from data collected from a variety of sources, including Github (Chen et al., 2021), Wikipedia (Schick et al., 2022), and others. These models may produce a wide variety of multimedia formats, such as audio, video, and photos, by utilizing this abundance of data. By examining and classifying the various Generative AI models that have been developed, our survey capitalizes on the breadth of this technology.

Groundbreaking Generative AI models include ChatGPT and DALL-E (Ramesh *et al*., 2022), which have catalyzed a new era in the synthesis and manipulation of digital content. The use and combination of these technologies have allowed for use cases throughout many different fields and applications which will be explored throughout this piece. Businesses' productivity has already been boosted by Generative AI, with research finding that 15% of tasks could be achieved significantly faster through the use of AI (Eloundou *et al*., 2023).

Deep neural networks, transformers, and other cutting-edge models like generative adversarial networks (Creswell et al., 2018) and variational autoencoders (Murphy, 2022) have been employed in recent years due to the ongoing increase in processing capacity. All of these models are skilled in simulating high-dimensional probability distributions of language or images from particular or broad domains because they can successfully represent the complexity of the data. Any input format, like text, can be converted into a variety of output formats, like video, by combining generative models with other methods that map the latent high-dimensional semantic space of language or images to multimedia representations of text, audio, or video. Generative models are useful in many applications because of their adaptability, which enables a smooth transition across multimedia formats. One of the most significant aspects of generative AI is its potential for

endless applications. From a variety of input formats, these models can be taught to produce really distinct multimedia representations, such as text, audio, or video. Generative AI, for example, can be used to develop video content from audio, realistic graphics from textual descriptions, or even musical compositions based on particular styles or emotions. Furthermore, by automating content generation and offering individualized experiences, generative AI has the ability to completely transform sectors like advertising, entertainment, and education. These models can save time and money while opening up new creative possibilities for both individuals and organizations because they can learn from a variety of data sources and produce a wide range of multimedia outputs. To sum up, generative AI models provide unmatched potential in content generation and transformation because of their access to vast amounts of data and intricate designs. They are essential tools in today's technologically advanced society because of their capacity to learn from multiple sources, create a wide range of multimedia forms, and transform inputs between different formats.

Surveys of LLMs and generative AI have been conducted in more recent work, discussing various uses of the technology (Zhao et al., 2023; Cao et al., 2023; Zhang et al., 2023a-c; Aydın and Karaarslan, 2023). Previous surveys have examined particular generative AI technology types, like text-to-image (Zhang et al., 2023a-c). Other studies, like the one conducted by Cao et al. (2023), have examined the effects of generative A on various businesses without delving deeply into the best models or technology to employ. Unlike previous surveys, this thorough overview attempts to provide a distinct viewpoint by showcasing not only the most well-known generative models and the technologies that underpin them, but also all of the various applications of this technology. We also provide a current competitive picture in this expanding business, along with the concepts that underpin this expansion. Text, pictures, video, 3D, code and software, speech, AI comprehension, business, gaming, music, biotechnology, the brain, multimodal, and computational efficiency are among the 13 categories covered by this site. A comprehensive taxonomy of the current technologies is provided in each section, including information on the models and tools that are available. The survey provides a methodical examination of these various AI applications, making it a vital resource for professionals, scholars, and researchers who wish to better understand the rapidly changing field of generative AI and its broad ramifications.

For instance, a 3D game designer can require different generative AI capabilities for a project. Under both 3D and gaming, he might locate a solution for his 3D AI wants, receiving more precise outcomes and distinct responses. He might also discover ways to expand his business. needs of his under text and business. We are certain that this poll will give users a clear picture of how generative AI is developing and where they can locate the technology they require.

In this article, we provide the idea of a comprehensive assessment focused on the most popular generative AI applications, which are significantly changing fields like business operations, design, and videogames (Liu et al., 2021; Kulkarni et al., 2023). The difficulty users have locating the produced programs in each unique application field supports the need for a thorough reference tool that classifies models.

In order to guarantee the study's inclusion and thoroughness, a great deal of information on generative Artificial Intelligence (AI) applications was gathered online. The inquiry entailed a thorough examination of a number of sources, including the official websites of businesses, where technical specifications and comprehensive documentation were examined. In order to gather useful information about practical uses, customer reviews, and updates pertaining to generative AI technologies, a review of businesses' social media accounts was also conducted. The goal of this two-pronged strategy was to record both official, formal data and dynamic, community-driven viewpoints on these programs.

The collected data was then subjected to a thorough categorization procedure. Distinct traits, functions, and underlying algorithms found during the evaluation of internet sources served as the foundation for the development of the classification system. Through an iterative method, each category was improved while remaining aware of new developments in the field of generative AI. In addition to enriching the dataset, the classification system offered a comprehensive perspective of the real-world applications and changing trends in the sector by combining data from many internet sources, such as official channels and community forums. This synthesis strengthens the study's general validity and applicability of the results offered in this scholarly investigation.

Basic Taxonomy of Models

This article examines the emerging uses of generative AI, emphasizing how it can revolutionize a variety of fields, including design, biology, business, and the arts. We accomplish this by breaking down generative AI into 13 components based on the output generated, the context in which it is utilized, and the technology's commercial applications. Given that the output is text, the reader might notice that a large number of models fall within the text category. Or that the text category might potentially include a lot of copy-writing models. The

A potential user of generative AI technologies can easily

ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 2, 2024

identify the technology they will be employing based on the use case by using the classification that is displayed below. The categories into which we have classified existing generative AI technologies are presented in this first section. We provide an overview of the many categories here: The goal of generative AI technologies in the text area is to produce and work with text in natural language. Among these technologies are language models, like OpenAI's GPT models, that can produce text that appears human. Although chatbots like Google's BARD and OpenAI's ChatGPT are the most well-known of these models, this category also includes additional models. These include chatbots, scientific language models, and text-writing helpers. The models' ability to generate text as an output was the primary requirement for this category. The production and modification of visual pictures is the main emphasis of generative AI systems in the images area. This category's primary requirement was that the output be an image. This can comprise image-editing models and image-creating models that can produce images from textual descriptions. To keep things simple, the category was separated into three categories: editing, realistic image production, and artistic image creation. Some models are randomly assigned to one of the categories if they completed two or more of these activities. Text-to-layouts and text-to-molecular representations are examples of additional models in this area that did not fit into the previously described categories. In the video sector, generative AI technologies are designed to produce and work with video material. The primary requirement for this category was that the finished product had to be a video. This primarily comprises models for creating new videos based on textual descriptions. Post-production, text-to-scene creation, text-to-motion capture, image-to-video, and video dubbing are other models. Generative AI technologies in the 3D category also operate with 3D, with an emphasis on creating and modifying three-dimensional environments and objects. Utilizing the output as a fully developed 3D model was the primary criterion. A 4-D model and a 3-D model created especially for the metaverse are also included. Text, a single image, pictures, and 2D models are examples of inputs. With an emphasis on code and software, generative AI technologies in this area seek to automate the code-writing and software-development processes. The primary requirement was that the finished product be a code. Text-to-code, text-to-websites, text-to-software, and text-to-apps are some of the many categories that fall under this. Other less common types are code translators, text-to-software, text-to-RPA, and designs-to-code. Adept, a firm that wants consumers to interact with computers using only text input, was the target market for the text-to-software category. It is placed here for this reason. In the speech category, generative AI systems concentrate on producing and modifying verbal communication. An input can be converted into a spoken output by any of these methods. This is separated into speech editing, speech-to-speech, and text-to-speech. In terms of AI comprehension, models that transform an input into a text output are referred to as generative AI technologies. The demand for a category that would compile models capable of converting a large number of inputs into voice led to the creation of this specific category. Speech, photos, music, video, metaphors, semi-structured and structured data, movies, and generating regions are all included in the inputs. In terms of business, generative AI technologies in the business category concentrate on using AI to enhance decision-making and business procedures. Businesses may find use for several of the models described in the previously mentioned categories, such as ChatGPT in the text category or Midjourney in the image area. People in general enterprises use this category to identify models for their operations in spite of this. These are separated into three categories: company operations, innovative business models, and marketing. Regarding gaming, the goal of generative AI technology in this field is to greatly simplify game development for developers. For their objectives, they make use of text, 3D, and image models. They are separated into characters and video game development. Regarding music, the development and modification of musical content is the main emphasis of generative AI technology in this field. This area covers dance-to-music models, musical editing, and music generation. Applying generative AI to biological research and medicinal applications is the goal of the biotechnology category's generative AI technologies. This can include drug discovery tools that can find novel drug candidates and models that can predict the structure of proteins or DNA sequences. Although the Business category was created due to the large number of generative AI applications in this domain, some of these models may have been put there. Regarding the human brain, the brain category's generative AI technologies concentrate on using generative AI to facilitate communication. Both brain-to-text and brain-to-image models fall under this category. A section was included for multimodal technologies to accommodate models that incorporated many other categories, including ChatGPT-4, which can generate both text and visual outputs. In order to accommodate Alphatensor, a revolutionary AI tool for algorithm discovery, we lastly add a category for computational efficiency (Kauers and Moosbauer, 2022).

It is expected that this methodical classification will be extremely valuable in the future. Researchers, academics, and professionals can better understand the changing environment of generative AI by using the survey, which provides an organized examination of various generative AI applications.

AI and its broad ramifications. The categorized

information, for example, may be useful to a 3D game designer looking for generative AI solutions for a project, yielding particular results in both the 3D and gaming categories. Likewise, both business and text categories offer alternatives for users with different business demands. In order to help users choose the best models for their particular projects in the future, this poll attempts to provide them a thorough grasp of how generative AI is developing.

Generative AI Applications

In this section, we will introduce a broad overview of generative AI applications divided into subsections according to every different topic.

Text

Since the introduction of ChatGPT, text models—particularly those focused on conversational chatbots—have completely transformed artificial intelligence. These models, which are supported by natural language processing and huge language models, have a number of extremely helpful features, including sentiment analysis, code generation, writing assistance, summarization, and language translation. Because of the capabilities of ChatGPT, an application that millions of users already use, they have been the main focus of generative AI. According to Noy and Zhang (2023), ChatGPT is believed to increase productivity by 40% in basic writing tasks, which is very promising for this technology. According to statistics, the output of GPT-2 cannot be distinguished from human-written material (Kreps et al., 2022). Technology has the potential to replace human output, at least in some types of writing, if the previous GPT model is unaffected by human output.

AI that can converse One of the most talked-about subjects in AI has been conversational AI. These services function as chatbots that can translate text inputs into text outputs for a wide range of functions. ChatGPT is the most well-known example, and it now contains built-in features to search the internet for information. Bing AI, Duck Assist, Google's BARD powered by LaMDA, ChatGPT's beta version, Metaphor, and Perplexity AI are some other models with updated data. Stanford Alpaca and Claude are two such chatbots that lack up-to-date knowledge.

Large Language Models, or LLMs, are the engine behind them. Transformer language models, such as GPT-3, Galactica, LLaMA, and PaLM, that include hundreds of billions (or more) of parameters and are trained on enormous amounts of text data are referred to as large language models (LLMs) (Zhao et al., 2023). These include mathematical reasoning, text production, programming support, spatial reasoning, commonsense reasoning, and more (Joublin et al., 2023; Tian et al., 2023; Frieder et al., 2024). Regarding business operations, there are numerous uses, including risk management, inventory optimization, and demand forecasting (Bahrini et al., 2023). Just as the capabilities of LLMs are being found, many of the capabilities are still being investigated at the time these articles were written.

Author-to-Text Simulation Recently, the models have demonstrated the ability to replicate specific writing styles. According to recent examples, LLMs can write as Lovecraft (Garrido-Merchán et al., 2023) or Daniel C. Dennett (Schwitzgebel et al., 2023). According to Dennett's paper, experts on his work were 51% successful in differentiating between the work of the philosopher and the broad language model. According to Lovecraft's paper, human readers who have never heard of Lovecraft before cannot tell the difference between texts published by ChatGPT and those written by the author. These are outstanding accomplishments that demonstrate how well language models may mimic writing through refinement.

Medical advice via text Through fine-tuning, big language models have also been shown to be helpful for initial medical recommendations. It is important to note that these models are still not entirely safe for this purpose and should not currently be used to replace humans. GlassAI, Med-PaLM 2, Chatdoctor, and your doctor AI are a few of these models (Singhal et al., 2023). They have demonstrated encouraging skills to obtain medical information, analyze it, and respond to inquiries about medicine in a manner similar to that of doctors. 86.5% was the highest score Med-PaLM 2 received on the MedQA dataset. Once more, these models demonstrate a remarkable ability to fine-tune and provide appropriate responses. Hippocratic AI (Hippocratic AI-hippocraticai.com, 2023), the largest startup in this field, has created LLMs that perform better than GPT-4 on medical datasets. This is particularly important right now because advancements have already been made in areas that have a direct bearing on people's quality of life.

Other uses of text-to-science can be observed in science, where Minerva (Lewkowycz et al., 2022) and Galactica (Taylor et al., 2022) have surfaced. Scientific language can be stored, combined, and reasoned about by Galactica, a big language model. Minerva is a large language model designed for college-level quantitative reasoning activities including science, math, and engineering difficulties. Despite their inability to completely replace human reasoning on these tasks, the models demonstrate encouraging outcomes in a very difficult area of research. Indeed, these models are recommended by current research.

Help with Writing Live writing support is another application for generative AI. Although previously stated chatbots like ChatGPT can be utilized for this, specialized programs like PEER (Schick et al., 2022) and GrammarlyGO (Grammarly, 2023) have been developed. Grammarly developed GrammarlyGO, a writing tool that can produce drafts, outlines, responses, and

changes. PEER is comparable to Grammarly's program, except it's optimized for scholarly articles and offers an explanation of how it works.

Images

Since the release of DALL-E 2 in 2022, image generative AI has continued to expand. This technology has been very helpful for both creative and professional uses in picture editing and in making images from written instructions. Regarding the production of art, it has been groundbreaking and has expanded the limits of creativity. Thanks to innovative apps like Midjourney, which have produced remarkably lifelike photos, photorealism seems to be drawing closer in the realm of image production.

Regarding creative pictures, there are a plethora of sites that provide text suggestions for the development of artistic images. One example is making pictures based on text instructions; two public sites where such models may be used are Mage. Space, Night Cafe, and DALL-E 2 are all examples of places that generate creative images by using steady diffusion. Neural and the smartphone app Wonder are two more platforms. Wonder allows users to create creative images. Love is an AI-driven platform for enhancing and manipulating music, video, and images; it includes an Art generator with a wide variety of genres to choose from, including Fantasy and Science Fiction. The platforms used by DALL-E and Midjourney for picture creation are different from these others.

There are further creative picture jobs where these models have shown usefulness. With the aid of Tattoos AI, you can create tattoos. With Supermeme AI, you can create memes. With profile image AI, you can design beautiful avatars using samples of yourself.

Generative AI for picture editing has been a boon in this area. Some practical applications include Alpaca AI, I2SB (Liu et al., 2023b), and Facet AI. These apps can do inpainting, outpainting, upscaling, super-resolution, deblurring, and depth map creation, among other things. Photoroom AI is a program that uses generative AI to edit photographs; it can remove objects and backgrounds from digital photos. A lot of people who aren't graphic designers can use natural language cues to alter their photos.

The Tencent Face Restoration tool demonstrates that even facial restoration may be accomplished using generative AI. One of the foundations of generative AI and deep learning, GANs allow them to do this. Stable diffusion reimagine is a tool for creatives that lets them make several versions of the same picture.

A great number of models have been developed that allow for the fabrication of realistic visuals. Some of them include PrimeProfile, Runway ML text-to-image, Craiyon, Bing AI Image Creator, Booth AI, Imagen, Midjourney, Muse, PrimeProfile, and stable diffusion ML (Li et al., 2023a–b), Yu et al., 2022, and Chang et al., 2023, respectively. They try to generate lifelike results using text inputs. Generative AI has several applications beyond text-to-image conversion. By analyzing picture samples, it can generate photorealistic visuals. You may use example photographs to generate headshots using Aragon AI, Avatar AI, and PrimeProfile.

Generative AI may enhance the design process. The use of latent diffusion to transform text into layouts is shown in Play (Cheng et al., 2023). Additionally, Autodraw is a model that can transform basic drawings into forms. You can speed up the design process using one of these apps.

Video

Storytelling is aided by video generative AI. The difficulty of video creation means that this discipline is still in its early stages of development. However, the use cases mentioned, such as digital human videos, video dubbing, and human motion capture, are groundbreaking and have the potential to greatly increase productivity.

Another use of image-to-video generative AI is in

Virtual reality applications greatly benefit from image-to-video generation. The GeoGPT model (Ren and Wang, 2022) and the SE3DS model (Koh et al., 2023) are two examples of generative AI models. Using a single scene picture and a trajectory of big camera moves, GeoGPT offered a new way to create a consistent long-term movie. Using new viewpoints—even those that expand far beyond the original pictures while preserving 3D consistency—SE3D generates high-resolution photographs and videos.

Models that can convert text to video are in their infancy. Meta Make a video, Imagen, Phenaki, and Runway Gen-2 are some of them. In order to generate visual outputs, Imagen video employs a cascade of diffusion models. The meta-research-created video-generation model known as "Meta make a video" can edit videos, convert images to videos, and convert text to videos. They have shown encouraging indicators and have been helpful with basic films, but they still have a long way to go before they can provide genuine results. Phenaki uses text prompts to generate films that are several minutes long. In addition, you may enter text, video, or images into Runway Gen-2, and it will make videos. By using a pre-trained text-to-image model, CogView2, CogVideo is able to produce shorter films in GIF format.

The use of digital people in video production is only one of many potential uses for these video models, which might ultimately result in cost savings. A variety of AI applications may be used to generate professional films using various avatars, such as Colossyan, Elai, Heygen, HourOne, Rephrase, and Synthesia. Several of these systems integrate this technology with 120+ languages to generate speech, like Synthesia.

Generative AI may also be used to turn articles into videos. With the use of Generative AI and an article as input, the smartphone software SuperCreator can make short movies for platforms like TikTok, Reels, and Shorts. Video synthesizers also turn articles into videos that may be shared on YouTube.

Businesses may greatly benefit from generative AI as it can lead to more personalized videos. Tavus AI, a platform for creating personalized films that automatically adjust the settings for each viewer, is a prime example of this. Additionally, D-ID generates real-time video using generative AI technologies to provide an immersive experience that mimics human behavior.

Artistic video generating is another area where they shine. One such software is Kaiber, which lets users make creative films by combining text and images. Even in the realm of filmmaking, there is Opus AI, a text-to-video generator that pays close attention to every detail, including scenes, characters, conversation, and visual effects.

Scenescape (Fridman et al., 2024) is a technique for text-driven perpetual view creation and human motion diffusion model; Riverside AI is an AI-powered video production site with edition capabilities. Other prominent approaches for video generating include these.

Because videos are multi-dimensional, this is an emerging area, as previously said. Regardless, this is not the end of the story; research has shown that conventional and synthetic movies teach the same material (Leiker et al., 2023). This demonstrates great potential for advancements in technology that might enable far quicker learning. Let it be known that instructional films are not the most intricate. However, sectors like tourism, education, and journalism might benefit from automation of basic video making, leading to increased productivity.

3D

A simple text prompt, picture, or video input is all that is needed to start 3D modeling. Many fields rely on them, including those dealing with 3D models, such as game development, the metaverse, and city planning. The automation of 3D modeling is an expensive endeavor for enterprises, but recent AI research suggests that it can create 3D models even with a poorly configured camera system (Pepe et al., 2023).

When it comes to picture-to-3D models, we may use either a single image or a collection of photographs to generate 3D models. A few well-known models that handle single-image inputs include GeNVS, Kaedim, MakeIt-3D, and Real Fusion (Chan et al., 2023; Tang et al., 2023a; Melas-Kyriazi et al., 2023). Scenedreamer (Chen et al., 2023), Neural-Lift-360 (Xu et al., 2023), NVIDIA Lion (Vahdat et al., 2022), and EVA3D (Hong et al., 2022) are all available for use with various picture inputs. In the case of humans in particular, there is PersoNeRF (Weng et al., 2023) that can take representative photos of humans and create a three-dimensional model of them. Another option is to use 2D pictures to create a 3D model. Additionally, we can use Deepmotion and Plask AI to convert video inputs into 3D models. Lastly, we can use NVIDIA LION to generate a 3D model from geometric coordinates (Vahdat et al., 2022).

The metaverse is where this technology is most obviously at work. As an example, Metaphysic AI enables artists to turn visual inputs into the production camera feed with minimum latency with generative AI. Versy AI is another company that has coupled Generative AI with the metaverse.

Adobe Firefly, dream fusion, GET3D, Magic3D, Synthesis AI, and Text2Room are among the most significant text-to-3D models with respect to text inputs (Gao et al., 2022; Lin et al., 2023). They used text inputs to generate textured 3D objects. Mirage is a 3D program that can make animated 3D objects from 3D models. Using a dynamic scene generator, MAV3D (Singer et al., 2023) demonstrates that generative AI can even produce 4D models.

Even while these technologies can't replace
present processes, the capabilities open the door to boundless opportunities. We are optimistic about the potential possibilities of this technology since Google is already using generative AI for Google Maps (Li et al., 2023c).

The Program and Its Source Code

Since its creation, GitHub Copilot and ChatGPT have been very helpful to developers. The user may create and design websites with the aid of these models using Natural Language. Additionally, they might be of assistance with the programmer's more routine duties, such writing documentation. The most audacious program, Adept, claims that generative AI may enable people to converse with computers only via language. One potential huge technical leap forward is the democratization of code, which will make these systems more easier to use for professionals who don't have a technical background. A recent research found that GitHub Copilot, the most popular generative AI tool for code, enhances productivity by 55% (Peng et al., 2023).

Converting Text into an App: Numerous Generative AI systems exist for the purpose of app generation. Some examples of generative AI platforms that allow users to quickly construct online applications using text prompts include Berri AI, Debuild AI, Flutterflow, Google Generative App Builder, Imagica AI, Scale Spellbook, Second AI, and ZBrain. Text and data inputs have made even the design of LLM apps accessible to professionals without technological expertise.

Microsoft Security Copilot is an AI-powered security analysis tool that can convert text into cybersecurity metrics, analyze signals, and determine risk exposure. It also allows for faster responses to attacks and improves cybersecurity using natural language.

Mintlify and Stenography, two text-to-coding documentation tools, have recently arisen as excellent methods for using generative AI in code documentation. Research has indicated that programmers spend around 5% of their time coding on documentation, therefore any

progress in automating this process would be much appreciated.

Generative AI has also enabled code translation from text to code; one of the most helpful tools in this regard is the Vercel AI code translator. This has the potential to greatly improve the efficiency with which businesses switch programming languages.

Coding text to Excel according to certain languages,

A lot of people have looked at using generative AI in Excel to generate spreadsheet code. Among these applications are Data Sheets GPT, Excel Formula Bot (Formula Bot AI Excel formula generator), Google Workspace AI-Sheets, and Sheets AI. AI Office Bot not only generates formulas rapidly through text prompts, but it also explains them. This is particularly useful because Excel is one of the most commonly used tools in office work, and workflows could potentially see improvements in productivity.

Text-to-multilingual coding there are several software for multilingual code generation through just text inputs. Although ChatGPT is widely used for coding, there are many more generative AI applications that are being created for that purpose. While most of them work as coding assistants, they are also able to generate code through text prompts. Some of them are Alphacode (Li et al., 2023a-c), Amazon Codewhisperer, BlackBox AI, code complete, CodeGeeX, Codeium, GitHub Copilot, GitHub Copilot X, GhostWriter Replit, Mutable AI (mutable.ai. They are used to complete, explain, transform, and generate code. They generate new lines based on context and syntax. As we can observe, it is one of the fields with the biggest amount of applications. They can be personalized to your writing style; Cod

Applications such as AI2SQL and search AI can convert text into SQL code.

Word-to-website translation for topics like Mutiny, Durable, and website development. Both programs take text input and output a webpage with text and photos. In particular, we have three apps—Diagram AI, Galileo AI, and Uizard AI—that use Generative AI to generate excellent user interfaces and optimize the consumer experience. Even more impressively, the.com automates the process of creating web pages, allowing businesses to build unique profiles for each client.

New technologies in the realm of computer programming have arisen, some of which are unrelated to generative AI code. Locofy is one example of a design-to-code technology that may be used to create code for online and mobile applications. In addition, text-to-automation technologies are available via Drafter AI, a platform that can automate complex analytical activities, and Lasso AI, which can automate any process using natural language. Efforts by Even Adept to enable text-based user interaction with all aspects of a user's computer have recently surfaced.

Public Speaking

Artificial speech models attempt to replicate human speech. The development of speeches has been made simpler by text-to-speech technology. The use of generative AI has made voice cloning a breeze with other speech-to-speech technologies. Podcasts, YouTube videos, and assisting mutes in communication are just a few of the many potential future uses for this technology.

The most significant models are VALL-E and speech-to-speech ones, such as ACE-VC (Hussain et al., 2023). VALL-E is able to transpose written words into speech with natural intonation and emotion based on the surrounding text using a three-second audio recording of a human voice.

In the same vein as the previously mentioned models, Supertone AI offers voice editing.

text-to-speech systems for the purpose of creating spoken language,

With the use of generative AI, it is now simple to record one's voice in response to text. Coqui, Descript Overdub, EllevenLabs Listnr, Lovo AI, Resemble AI, Replica Studios, Voicemod, and Wellsaid are just a few of the many platforms that have been developed. The most crucial model is AudioLM, which is Google's framework for consistently producing high-quality audio (Borsos et al., 2023) and is widely used. The Video Recordings-to-voice Dubverse is an excellent tool for video dubbing projects as it converts video recordings into voice.

Deep Learning for AI

The ability of AI to convert many forms of media into human-readable text, video, audio, and more has advanced to a satisfactory degree. This serves a significant purpose due to

machine learning's capacity to understand human speech and simplify it into simpler text. We can build more robust AI models if we can convert any input into text, which will make it easier to grasp and even use as input in other technologies. As an example, AI achieves an average inaccuracy of 1.15 percent in speech-to-text tasks, which is quite near to human performance (Radford et al., 2023). Computers will soon be able to allow rapid comprehension of huge and complicated datasets if they can transform all inputs into text, which is why these discoveries are quite encouraging. This area has enormous potential for the future of report generation, since it may include audio, graphics, and text into a written output that consumers will find very useful.

Modeling of generating regions into text has also made use of this technique. An object understanding transformer called GriT (Wu et al., 2022) uses pairings of regions and texts to find and characterize things. Object detection jobs may benefit from this.

There are other technologies that can convert photos to text as well. A number of domains may benefit from these technologies, including computer vision, which can aid AI in its comprehension of user-generated material. Alayrac et

al. (2022), Kirillov et al. (2023), and VisualGPT (Chen et al., 2022) are a few examples of applications that include these technologies.

Various inputs-to-text Some systems even allow us to convert various inputs into text. Primer AI can process and respond to large volumes of media in real time, including text, photos, audio, and video. This information may be better understood and used to safeguard democracy and national security. With the use of speech AI, research and marketing teams may transcribe and analyze unstructured audio, video, and text to get competitive insights. Generative AI's ability to swiftly evaluate large and unstructured data sets is shown by both technologies. Primer can help us comprehend and act on it, and Speak AI can provide us insights rapidly.

Natural language is another tool that pdf-to-text generative AI may utilize to extract data from documents. The two programs are MapDeduce and ChatDOC. Using natural language searches, they may swiftly find what they need in PDFs and summarize it. This functionality has also been included into ChatGPT's model from late October 2023.

One of the most prominent areas in speech-to-text has been speech-

transcribing and subtitle technology are really beneficial. The following applications are included: Cogram AI, Deepgram AI, Dialpad AI, Fathom Video, Fireflies AI, GoogleUSM (Zhang et al., 2023a-e), Papercup, Reduct Video, Whisper (Radford et al., 2023), and Zoom IQ. It's important to note that not all of these technologies only execute speech-to-text jobs; in fact, some of them do many more. Using Deepgram AI, we can determine who is speaking,

syntax, as well as the search terms. The artificial intelligence features of Dialpad automate client touchpoints, provide real-time suggestions, and summarize calls. In addition to translating, Papercup can also generate voices that seem natural. The last point is that zoom has made use of AI in its systems, which includes tools like chat summaries and email composition. Workflow optimization may be seen by merging several generative AI technologies.

Another area where table data-to-text generative AI has proven effective is in turning data tables into text. Defog AI, MURMUR (Saha et al., 2022), and TabT5 are a few examples of Generative AI apps that serve this goal. Unstructured data may be understood by Murmur in particular. A significant improvement in the efficiency of company decision-making based on the rapid interpretation of table data is possible if we succeed in perfecting this technique.

This is something that videos-to-text flamingo can even handle when given video inputs. Both TwelveLabs and Minotaur were identified as video inputs (Goyal et al., 2023). TwelveLabs takes a video input and turns it into vector representations of important elements including objects, actions, voice, screen text, and people. You can do a fast search using these vectors. In its long-form films, Minotaur takes on query-based video comprehension. As it simulates the proper identification of visual scenes in movies, another model dubbed MOVIECLIP (Bose et al., 2023) was discovered to be particularly beneficial in this field. This innovation allows us to see the beginnings of computers' ability to comprehend large, unstructured datasets.

Business

There is no doubt that generative AI will have far-reaching effects on the corporate world, making use of several of the aforementioned technologies in the process. Automating more creative and expensive operations, including creating marketing materials, PowerPoint presentations, or decreasing repetitive chores, may help firms save money. It has the potential to give rise to whole new kinds of AI-powered companies, like Truewind, which automates accounting, or Harvey, which automates litigation. We can only speculate as to the extent to which generative AI may impact business operations in the ways outlined below, given its relative youth. Goldman Sachs estimates that generative AI has the potential to propel GDP growth by 7%, highlighting the significance and influence of this technology.

Other sectors have also felt the effects, according to accounting records. Accounting is one area that has adopted Truewind, an AI-powered bookkeeper that aims to improve accuracy and transparency.

Generative AI may also be used to produce advertisements, as shown by several applications as Omneky, Pencil, Waymark, Clickable, and Ad Creative AI. Waymark, the final one, is great because it searches the web for local content and then uses that data to create films.

commercial information. Additionally, LensAI is helpful since it allows for more precise ad targeting by recognizing and pairing relevant advertising with specific items, logos, activities, and context. Generative AI may also support the storytelling in these commercials with apps like Subtxt and AI 21 Labs.

Another field that has made use of generative AI is architecture, namely with tools like Autodesk Space creator and SWAPP AI. SWAPP generates comprehensive, precise, and correct architectural construction documentation and building information modeling (BIM) files by using smart, sophisticated algorithms. Architects, city planners, and property developers may create first-rate site plans with the help of Autodesk Spacemaker, an AI program hosted in the cloud. Actually, Zuma—AI-powered real estate assistance that automates lead generation—shows that generative AI may also aid in the real estate aspect of the process.

From a research perspective, business visualizations are currently available in apps that allow you to ask basic

ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 2, 2024

questions using natural language and get responses supported by actual data. These responses are presented in a visual format that facilitates process integration and expedites market search. The following are examples of such businesses: Alphawatch, Dataherald, Maya, and OpenAxis AI.

In order to harness the potential of generative AI, the organization creates visually appealing content. As shown with Microsoft Designer, which generates invites, digital postcards, images, and more, designs may be swiftly made with just text prompts. One example of generative AI in action is Brandmark, which generates logos and other business-related material like business cards. Another example is Looka AI. Namelix, brandinition, and brandsnap are three resources that might help you generate ideas for company names.

Another use of customer assistance generative AI is the automation of customer communication. One Reach AI, OpenSight AI, Brainfish, and Yuma AI are some of the applications that may help your organization create customized chatbots. Generative AI also has tools like InboxPro, Lavender, Smartwriter, and Twain that can automate e-mails. Email analytics and social media data are two examples of the kinds of technology that might help improve business processes. There are also voice-assistant systems like Poly AI.

When it comes down to it, Generative

AI via Sonantic, an editing and directing tool for text-to-voice acting.

Additionally, educational institutions have been impacted. This is mostly due to the introduction of ChatGPT into students' assignments (Li et al., 2023a-c) and the rise of firms like Broadn, which use language models and generative AI to enable individuals to design personalized learning courses that align with their individual learning styles.

With the help of feedback-generative AI tools like Venturus AI and Mixo AI, we may get instantaneous feedback on our business ideas.

When it comes to the tedious process in the finance industry, generative AI can be a huge help. BloombergGPT is a finance-focused large-scale language model developed from the ground up. Sentiment analysis, news categorization, named entity identification, and question answering are just a few of its many potential applications. We can see how this may be very useful for anyone working in the financial sector. When it comes to modeling, Quilt Labs AI is a platform that uses AI to turn financial data into models. One sector that could benefit greatly from the automation that generative AI might provide is the financial sector.

Apps for internal communications, such as Albus ChatGPT, ChatGPT for Slack, and Moveworks, which automate employee assistance and provide discussion summaries, may greatly aid in staff management.

An excellent illustration of an industry that has been profoundly impacted by Generative AI is the legal documentation sector. Some examples of apps and businesses in this space are Spellbook Legal, Harvey AI, Casetext CoCounsel, and Darrow AI. Contract analysis, due diligence, litigation, and regulatory compliance are all areas where Harvey AI may be of assistance by providing data-driven insights, suggestions, and forecasts. Darrow AI, on the other hand, may source cases for your legal business and do due diligence on them. In regards to Spellbook Legal, it evaluates and proposes the contract conditions using GPT-4. It is clear that Generative AI is taking over several tasks in the legal industry that might be mechanized with ease. An software called TaxGPT uses OpenAI's GPT to automate the process of filling out tax forms in this area.

Since generative AI may simplify the process of creating innovative regions and images, it has had a significant impact on marketing jobs. Anyword, Copy AI, Google Workspace-Gmail and Docs, Hyperwrite, Jasper, Letterdrop, Regie AI, Simplified AI, Type AI, and Writesonic are just a few of the many copywriting programs that have previously been created. Emails, website material, drafts, answers, marketing content, and product descriptions are all within the scope of capabilities. Many companies might benefit much from streamlining these procedures, as is readily apparent. To go the extra mile for your firm, Regie AI may modify the LLM's tone to match your brand's voice. Jasper can compose reports, emails, blogs, and social media postings; this is just one more example of how companies may enhance their operations by combining several generative AI technologies.

Companies like LA LA LAND, which uses an AI-powered digital model studio to display your 3D concepts as realistic models, are demonstrating how Generative AI is already impacting the modeling industry.

Through Salesforce's EinsteinGPT platform, which generates tailored content for all Salesforce clouds, we can see how firms are spending money on artificial intelligence (AI). All interactions with sales, service, marketing, commerce, and IT will produce content, revolutionizing the customer experience. A comprehensive program that streamlines all processes is AI Intern IO. It provides AI for a wide range of company activities, including text, reports, code, marketing, HR papers, legal documents, translations, and documentation.

Generative AI applications, like as Cohere AI, which provides LLM for text retrieval, generation, and classification, may help enhance product design and product production.

Automating routine processes is another area where generative AI may be useful to businesses. Bardeen AI, Magical AI, and Notion AI are just a few of the numerous programs that can do this. programs like these are tailored to repetitive operations and are particularly helpful for businesses who are looking to automate relatively straightforward procedures using machine learning.

Thanks to the abundance of pre-existing applications,

generative AI can boost sales. Applications that optimize contact centers and improve the customer experience include Cresta, Forethought AI, Grain AI, and Replicant. When it comes to customer care, Replicant is up to the task. Cresta and Grain are among the others that provide live support to call centers. Automatic note-taking, record-keeping, and insight-capture for client talks are automated by Grain AI, while Cresta turns real-time insights into real-time actions. When it comes to Forethought, their goal is to streamline the customer service process. In order to ensure that every sales call is well-prepared, the app Tennr was developed to provide ideal meeting prep materials. Copy Monkey AI is a software that can optimize your product's listing on Amazon and boost its organic ranking.

Similarly, it may facilitate laborious procedures used in scientific research. Applications like Koala, Elicit AI, ConsensusNLP, Agolo AI, and ArxivGPT demonstrate this. Finding articles, extracting important statements, and highlighting insights are some of the features of these programs. In particular, two chatbots tailored to the needs of scientists are Koala and ConsensusNLP. The use of generative AI for fact-checking has also been investigated by Golden.

Generative AI may help simplify the analyst's procedures by generating slides. Assisting with both slide creation and market research allows for this to be accomplished. Several applications exist that can generate presentations using natural language for the purpose of slide creation. Automatic Slides, Canva Docs to Decks, ChatBA, Decktopus, Gamma, Google Docs to Slides, Tome, and Slide AI are just a few of them.

Some of them, like Tome AI, just need a little textual instruction, while others, like Canva Docs, which turns papers into slide presentations, require lengthy texts to function. In addition, Decktopus can make valuable slide notes for you.

Apps like Clips AI, Pictory AI, Predis AI, Tweethunter, and Tweetmonk are designed to help in the development of social media content. Using AI, Clips AI and Pictory AI can transform lengthy pieces of material into shareable social media posts. Using your brand's voice, Predis AI creates visual and video content. Using either Tweethunter or Tweetmonk, you may get your brand's content tweeted. We can watch as Generative AI learns your brand's ins and outs and automates these tasks in no time. Through Bytepods, businesses may also take use of Generative AI to create their own podcasts.

Generative AI may also be useful for higher-level, more strategic divisions of a firm when making strategic decisions. Using GPT, applications such as Rationale AI may facilitate the development of many business assessments.

One last application of generative AI is to generate synthetic data that mimics real-world conditions for testing purposes. Hazy, Mostly AI, Octopize, and Tonic are some of the websites that may help you do this.

Additional features for trip planning include the ability to create itineraries; some examples of such applications are Roam Around, TripNotes, and the Kayak plug-in for ChatGPT. The capacity to build a visiting itinerary is shown by the first two, while the ability to search for hotels, flights, and more using natural language is demonstrated by the Kayak plugin.

Video game

The ability to create generative AI from text, 3D models, and photos will be a huge boon to the game industry. Text models are useful for narrative and character development, whereas 3D models aid in the creative process. We can learn a lot about the potential of Generative AI in other industries by looking at the gaming industry as a case study. Since research indicates that the gaming industry would experience more realistic gaming environments and more engaging and individualized video games as a result of using Generative AI, these developments seem to be quite noteworthy (Wu et al., 2023). Your patience is much appreciated. The rearranged data is shown here.

When making video games, generative AI may be a useful tool. This is shown by means of programs such as Latitude, Illiad AI, and CSM. In particular, PixelVibe facilitates the production of game assets by use of generative AI.

Characters from video games: We've discovered Character AI, ConvAI, in-world AI, and RCT AI Chaos Box. Using natural language, ConvAI and World AI create characters. You can get entire characters only by adding character settings.

In terms of RCT AI Chaos Box, it is an engine that utilizes Deep Reinforcement Learning to assess player inputs in real-time and produce NPC replies and new narratives on the fly.

One game that uses LLMs for open-ended text-to-level creation is Levels MarioGPT (Sudhakaran et al., 2024).

Armorlab is a piece of software that uses artificial intelligence to fuel the creation of textures for video games.

Music

Generative AI may also significantly enhance the process of creating music. Inputting a simple text prompt or music tracks as input are both viable options for accomplishing this aim. Data from some experiments suggests that people can't tell the difference between music created by humans and music created by artificial intelligence. A song produced by Drake and The Weeknd using artificial intelligence had over 600,000 listens on Spotify in only one day, proving this point. Due to the Intellectual Property concerns that came from exploiting someone else's music samples for your own profit, it only lasted 24 hours. Here we see firsthand how issues related to the usage of publicly available data might emerge from irresponsible use of generative AI. There is a vast diversity

of application cases for this technology that may enhance musicians' creativity beyond just making tunes. These tools may help artists overcome creative blocks by providing them with written suggestions, audio samples, and gestures to help spark ideas.

There have been several uses for text-to-music, which involves generating music using natural language instructions. The following systems are capable of producing music using basic natural language: Aiva, ERNIE-music (Zhu et al., 2023), Harmonai, Infinite Album, Jukebox, Mubert, Musico, Noise2Music (Huang et al., 2023), Sonify, soundful, and Splash AI. Based on language prompts, Splash AI may explore and mix samples, or it can use a sophisticated generative model trained on this dataset to generate original, high-quality music. Hyperpop, electronic dance music (EDM), glitch, phonk, trap, lo-fi, and hip-hop are just a few of the genres that users may make music in. The outcomes are remarkable even if the songs are just fifty seconds long. While the outcomes may vary, all of the models are comparable to Splash AI in that they generate musical compositions in response to a user-supplied natural language query.

Another kind of inputs-to-music model incorporates several kinds of

their technological advancements. In order to generate musical scores, musico responds to a variety of inputs, including motions, codes, and other noises. A new model called EDGE is making it possible to turn even dancing into music.

Applications like Moises AI and SingSong, which use Generative AI, may likewise enable musical editing (Donahue et al., 2023). The AI from Moises is useful.

practitioners of musical editing by removing vocals and altering tempo and timbre.

The field of biotechnology

Generative artificial intelligence tools facilitate the process of molecular modeling, which is useful in biotechnology. The fields of drug discovery and protein modeling may both benefit from this. Improvements in biotechnology may also be possible as a result of these technologies' continued development. The NASDAQ-listed Absci Corporation already employs generative AI in the development of new pharmaceuticals. Generative AI can rapidly generate realistic structures, increasing success rates to 20% (Winnifrith et al., 2023), which is a significant improvement over the relatively low success rates in protein modeling (Syrlybaeva and Strauch, 2023) for example. The potential for these technologies to significantly accelerate processes is clear.

Using NVIDIA Bionemo, a cloud service for generative AI in drug discovery, researchers have access to large-scale biomolecular AI models that can generate and forecast new compounds. Atomic AI, BigHat AI, Exscientia, Menten AI, and ProteinQure are just a few of the many firms that use Generative AI for medication innovation.

They develop antibody treatments for patients by using synthetic biology and machine learning, with BigHat AI serving as an example of their skills in the domain of drug discovery. Through continual improvement, AI models enhance molecules.

Hundreds of distinct compounds may be designed simultaneously using their own machine-learning methods. The next step is to measure the molecular characteristics and add them to the model. Pharmaceutical firms achieve this level of efficiency via the use of Generative AI in conjunction with testing feedback. In addition to producing high-quality data, they both boost efficiency. It established a partnership with Merck to work together on up to three drug development projects in November 2022.

As for protein modeling, several models that have been discovered include BARTSmiles, a generative language model for molecular representation (Chilingaryan et al., 2022), and Alphafold, a computer program that predicts protein structures for the full human genome (Ren et al., 2023). Protein design using Generative AI is at the heart of two newly discovered firms, Cradle and Profluent.

To see an example of AI in action in Protein

concept, we may examine Cradle's commercial activities. To forecast the structure of a single sequence or a multimer, Cradle employs Alphafold's model. Utilizing models based on extensive datasets, Cradle enhances the thermostability of your proteins. Generative AI allows you to construct your own proteins, much as BigHat AI.

foretell how things will go and see the outcomes graphically. Once you have finished, you may share your findings with the firm to help them refine their models.

Brain

A brain model is a piece of software that can read and interpret brain waves. They have the potential to help us decipher the significance of a person's brain waves. Helping the deaf and hard of hearing communicate with generative AI is one of their most exciting potential applications. Even though it's still in its early stages of development, the field is seeing encouraging outcomes; for example, Meta's models achieved 41% accuracy. These findings hold great promise for altering the way doctors see brains. In addition, generative AI has the potential to improve classroom instruction. Technological advancements have been shown to facilitate education in recent studies (Darmawan, 2012). Researchers will have a better understanding of brain communication and be able to implement more efficient forms of communication if these technologies enable them to translate brain signals into text.

Using a Diffusion Model (DM) known as stable diffusion, a novel approach termed brain signals-to-images Using Stable Diffusion for Brain Images (Tang et al., 2023b) may convert brain waves into pictures. The participants'

ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 2, 2024

previously observed photographs might be recreated using this model. While the findings are excellent, it would take a lot of resources to retrain the model and do the experiment again. The use of steady diffusion to reconstruct brain pictures has never been done before.
When looking for models that can convert brain impulses into text, we came across Meta AI's brain speech and Non-Invasive Brain Recordings (Défossez et al., 2023). Both of them are attempting to decipher speech using non-invasive brain recordings.

Different kinds of
The models may integrate many technologies from the list and make the most of them. All of the apps on this list accept various inputs, which may speed up the development of AI. The future of Generative AI may lie in initiatives involving multi-tasking agents like GATO. While some models, like text-to-slides, do make use of several generative AI technologies, they were chosen for this reason alone.
A multimodal chatbot is one that can take in text, images, and audio as inputs and then output text and images. GPT-4 is the latest iteration of this technology. The sector is seeing a major shift as the most proficient chatbot transitions to multimodal. Chatbots' capabilities and the types of data they may assist users with are both enhanced when they accept several inputs at once. Increasing input types

have the potential to increase the technology's independence by allowing for the integration of AI into more operations than previously possible.
Kosmos-1, described by Huang et al. (2024), is one multimodal language model that has several capabilities. The ability to comprehend and generate new languages is one sort of capability. Perception-language activities include tasks like visual question answering, multimodal discourse, picture captioning, and image identification with descriptions. Prismer, as described in Liu et al. (2023a), is a paradigm for vision language that makes use of multi-modal experts. Question answering, object identification, segmentation, and picture captioning are a few examples of tasks. While using training data that is up to two orders of magnitude less, our model is comparable with state-of-the-art vision models that are currently available. The PaLM-E model is an embodied multimodal language model (Driess et al., 2023). First, PaLM-E's primary goal in development was to serve as a model for robotics; it accomplishes this goal by solving a wide range of tasks using a wide range of robot kinds and modalities, including pictures, robot states, and neural scene representations. Simultaneously, PaLM-E is an adequately competent vision-and-language model. It excels in both visual and verbal tasks, including object detection, scene classification, and quoting poetry, as well as solving mathematical problems and generating code. In the world of multi-format input chatbots, the ERNIE bot stands out. It can take a text prompt and turn it into text, graphics, audio, and video. What's more, it can even offer speech in many local dialects of Chinese, including Sichuan. Initial expectations were not met by Ernie's user turnout. Still, the CEO has asserted that the chatbot can hold its own against ChatGPT.
The NVIDIA Picasso cloud service allows users to create and launch 3D, video, and picture apps that are driven by generative AI. It is part of a larger framework of integrated AI models for multimodal cloud services in generative AI. It combines text with images, video, and 3D models.
In fact, a framework known as HuggingGPT exists.
(Shen et al., 2024) that uses LLMs (see: ChatGPT) to link different AI models in ML communities (see: embracing face) in order to do AI tasks. With the use of language as a general interface, LLMs take on the role of controllers, managing existing AI models to complete complex AI tasks. Its remarkable performance in language, vision, and voice, among other difficult tasks, opens the door to more sophisticated AI. Using the same framework to combine many models may provide remarkable outcomes. Images with writing on them may be used to create sounds. The potential for AI to simultaneously optimize several industrial processes is almost limitless when these models are combined. An improvement in graphic design's efficiency might be achieved rapidly by simply merging text and picture models.
When it comes to Adobe FireFly, it's a family of models that takes text as input and outputs various media types such as photos, vectors, videos, and 3D models. It uses language commands to trigger the generation of pictures and other materials using many AI models. With Photoshop's new generative fill tool, users can easily add, expand, or delete material from photos using simple text prompts. Quickly removing people, adding puddles, and animals are just a few examples. It is also possible to make several drawings out of only one. Faster processes are the result of integrating all these functionalities into a single software.

When it comes to efforts to create generalist agents, GATO stands alone as an agent that goes beyond only text outputs. A generalist strategy that employs several modes, tasks, and embodiments is followed. Everyone on the same network can do things like push buttons, communicate, and play games simultaneously. Concerning Imbue (formerly generally intelligent), it is a business responsible for creating agents that are generally competent. They want to imbue machines with intellect and human values; they're worth $1 billion. Their goal is to implement AI systems that are matched with human levels, so that they can aid in scientific research and generalize to a variety of economically beneficial jobs. They do this by creating internal agent prototypes utilizing their large pre-trained

ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 2, 2024

models that are now focused on coding chores. This is the only firm that we are aware of that deals with AI agents. We also think their effort is important since Imbue is trying to make models that are easier to understand.

Artificial general intelligence (AGI) advancements are crucial because they bring computers closer to mimicking human intellect (Latif et al., 2023). By responding independently to human commands, AutoGPT became the first method for artificial general intelligence (Talebirad and Nadiri, 2023) in the field. As a result, the model's star rating on GitHub soared to over 150,000.

Computational Efficiency

Alphatensor Fawzi *et al*. (2022) is the first-ever Artificial Intelligence (AI) system for developing new, effective, and indubitably correct algorithms for essential operations like matrix multiplication based on reinforcement learning. The task given to Alphatensor was to improve the efficiency of matrix multiplications, which occur in many fundamental computations. Automating the algorithm discovery procedure is intricate, as the space of possible algorithms is enormous. That is why this model uses AlphaTensor, which is trained to play a single-player game where the objective is finding tensor decompositions within a finite factor space. AlphaTensor discovered algorithms that outperform the state-of-the-art complexity for many matrix sizes. Matrix multiplication is crucial for many areas from neural networks to scientific computing routines thus this study will surely have an effect on computing efficiency.

Conclusion

Since the release of DALL-E 2 in 2022, image generative AI has continued to expand. This technology has been very helpful for both creative and professional uses in picture editing and in making images from written instructions. Regarding the production of art, it has been groundbreaking and has expanded the limits of creativity. Thanks to innovative apps like Midjourney, which have produced remarkably lifelike photos, photorealism seems to be drawing closer in the realm of image production.

Regarding creative pictures, there are a plethora of sites that provide text suggestions for the development of artistic images. One example is making pictures based on text instructions; two public sites where such models may be used are Mage. Space, Night Cafe, and DALL-E 2 are all examples of places that generate creative images by using steady diffusion. Neural and the smartphone app Wonder are two more platforms. Wonder allows users to create creative images. Love is an AI-driven platform for enhancing and manipulating music, video, and images; it includes an Art generator with a wide variety of genres to choose from, including Fantasy and Science Fiction. The platforms used by DALL-E and Midjourney for picture creation are different from these others.

There are further creative picture jobs where these models have shown usefulness. With the aid of Tattoos AI, you can create tattoos. With Supermeme AI, you can create memes. With profile image AI, you can design beautiful avatars using samples of yourself.

Generative AI for picture editing has been a boon in this area. Some practical applications include Alpaca AI, I2SB (Liu et al., 2023b), and Facet AI. These apps can do inpainting, outpainting, upscaling, super-resolution, deblurring, and depth map creation, among other things. Photoroom AI is a program that uses generative AI to edit photographs; it can remove objects and backgrounds from digital photos. A lot of people who aren't graphic designers can use natural language cues to alter their photos.

The Tencent Face Restoration tool demonstrates that even facial restoration may be accomplished using generative AI. One of the foundations of generative AI and deep learning, GANs allow them to do this. Stable diffusion reimagine is a tool for creatives that lets them make several versions of the same picture.

A great number of models have been developed that allow for the fabrication of realistic visuals. Some of them include PrimeProfile, Runway ML text-to-image, Craiyon, Bing AI Image Creator, Booth AI, Imagen, Midjourney, Muse, PrimeProfile, and stable diffusion ML (Li et al., 2023a–b), Yu et al., 2022, and Chang et al., 2023, respectively. They try to generate lifelike results using text inputs. Generative AI has several applications beyond text-to-image conversion. By analyzing picture samples, it can generate photorealistic visuals. You may use example photographs to generate headshots using Aragon AI, Avatar AI, and PrimeProfile.

Generative AI may enhance the design process. The use of latent diffusion to transform text into layouts is shown in Play (Cheng et al., 2023). Additionally, Autodraw is a model that can transform basic drawings into forms. You can speed up the design process using one of these apps.

Video

Storytelling is aided by video generative AI. The difficulty of video creation means that this discipline is still in its early stages of development. However, the use cases mentioned, such as digital human videos, video dubbing, and human motion capture, are groundbreaking and have the potential to greatly increase productivity.

Another use of image-to-video generative AI is in

Virtual reality applications greatly benefit from image-to-video generation. The GeoGPT model (Ren and Wang, 2022) and the SE3DS model (Koh et al., 2023) are two examples of generative AI models. Using a single scene picture and a trajectory of big camera moves, GeoGPT offered a new way to create a consistent long-term movie. Using new viewpoints—even those that expand far beyond the original pictures while preserving 3D consistency—SE3D generates high-resolution photographs and videos.

Models that can convert text to video are in their infancy. Meta Make a video, Imagen, Phenaki, and Runway Gen-2 are some of them. In order to generate visual outputs, Imagen video employs a cascade of diffusion models. The meta-research-created video-generation model known as "Meta make a video" can edit videos, convert images to videos, and convert text to videos. They have shown encouraging indicators and have been helpful with basic films, but they still have a long way to go before they can provide genuine results. Phenaki uses text prompts to generate films that are several minutes long. In addition, you may enter text, video, or images into Runway Gen-2, and it will make videos. By using a pre-trained text-to-image model, CogView2, CogVideo is able to produce shorter films in GIF format.

The use of digital people in video production is only one of many potential uses for these video models, which might ultimately result in cost savings. A variety of AI applications may be used to generate professional films using various avatars, such as Colossyan, Elai, Heygen, HourOne, Rephrase, and Synthesia. Several of these systems integrate this technology with 120+ languages to generate speech, like Synthesia.

Generative AI may also be used to turn articles into videos. With the use of Generative AI and an article as input, the smartphone software SuperCreator can make short movies for platforms like TikTok, Reels, and Shorts. Video synthesizers also turn articles into videos that may be shared on YouTube.

Businesses may greatly benefit from generative AI as it can lead to more personalized videos. Tavus AI, a platform for creating personalized films that automatically adjust the settings for each viewer, is a prime example of this. Additionally, D-ID generates real-time video using generative AI technologies to provide an immersive experience that mimics human behavior.

Artistic video generating is another area where they shine. One such software is Kaiber, which lets users make creative films by combining text and images. Even in the realm of filmmaking, there is Opus AI, a text-to-video generator that pays close attention to every detail, including scenes, characters, conversation, and visual effects.

Scenescape (Fridman et al., 2024) is a technique for text-driven perpetual view creation and human motion diffusion model; Riverside AI is an AI-powered video production site with edition capabilities. Other prominent approaches for video generating include these.

Because videos are multi-dimensional, this is an emerging area, as previously said. Regardless, this is not the end of the story; research has shown that conventional and synthetic movies teach the same material (Leiker et al., 2023). This demonstrates great potential for advancements in technology that might enable far quicker learning. Let it be known that instructional films are not the most intricate. However, sectors like tourism, education, and journalism might benefit from automation of basic video making, leading to increased productivity.

3D

A simple text prompt, picture, or video input is all that is needed to start 3D modeling. Many fields rely on them, including those dealing with 3D models, such as game development, the metaverse, and city planning. The automation of 3D modeling is an expensive endeavor for enterprises, but recent AI research suggests that it can create 3D models even with a poorly configured camera system (Pepe et al., 2023).

When it comes to picture-to-3D models, we may use either a single image or a collection of photographs to generate 3D models. A few well-known models that handle single-image inputs include GeNVS, Kaedim, MakeIt-3D, and Real Fusion (Chan et al., 2023; Tang et al., 2023a; Melas-Kyriazi et al., 2023). Scenedreamer (Chen et al., 2023), Neural-Lift-360 (Xu et al., 2023), NVIDIA Lion (Vahdat et al., 2022), and EVA3D (Hong et al., 2022) are all available for use with various picture inputs. In the case of humans in particular, there is PersoNeRF (Weng et al., 2023) that can take representative photos of humans and create a three-dimensional model of them. Another option is to use 2D pictures to create a 3D model. Additionally, we can use Deepmotion and Plask AI to convert video inputs into 3D models. Lastly, we can use NVIDIA LION to generate a 3D model from geometric coordinates (Vahdat et al., 2022).

The metaverse is where this technology is most obviously at work. As an example, Metaphysic AI enables artists to turn visual inputs into the production camera feed with minimum latency with generative AI. Versy AI is another company that has coupled Generative AI with the

metaverse.
Adobe Firefly, dream fusion, GET3D, Magic3D, Synthesis AI, and Text2Room are among the most significant text-to-3D models with respect to text inputs (Gao et al., 2022; Lin et al., 2023). They used text inputs to generate textured 3D objects. Mirage is a 3D program that can make animated 3D objects from 3D models. Using a dynamic scene generator, MAV3D (Singer et al., 2023) demonstrates that generative AI can even produce 4D models.
Even while these technologies can't replace
present processes, the capabilities open the door to boundless opportunities. We are optimistic about the potential possibilities of this technology since Google is already using generative AI for Google Maps (Li et al., 2023c).

The Program and Its Source Code
Since its creation, GitHub Copilot and ChatGPT have been very helpful to developers. The user may create and design websites with the aid of these models using Natural Language. Additionally, they might be of assistance with the programmer's more routine duties, such writing documentation. The most audacious program, Adept, claims that generative AI may enable people to converse with computers only via language. One potential huge technical leap forward is the democratization of code, which will make these systems more easier to use for professionals who don't have a technical background. A recent research found that GitHub Copilot, the most popular generative AI tool for code, enhances productivity by 55% (Peng et al., 2023).
Converting Text into an App: Numerous Generative AI systems exist for the purpose of app generation. Some examples of generative AI platforms that allow users to quickly construct online applications using text prompts include Berri AI, Debuild AI, Flutterflow, Google Generative App Builder, Imagica AI, Scale Spellbook, Second AI, and ZBrain. Text and data inputs have made even the design of LLM apps accessible to professionals without technological expertise.
Microsoft Security Copilot is an AI-powered security analysis tool that can convert text into cybersecurity metrics, analyze signals, and determine risk exposure. It also allows for faster responses to attacks and improves cybersecurity using natural language.
Mintlify and Stenography, two text-to-coding documentation tools, have recently arisen as excellent methods for using generative AI in code documentation. Research has indicated that programmers spend around 5% of their time coding on documentation, therefore any progress in automating this process would be much appreciated.
Generative AI has also enabled code translation from text to code; one of the most helpful tools in this regard is the Vercel AI code translator. This has the potential to greatly improve the efficiency with which businesses switch programming languages.
Coding text to Excel according to certain languages,
A lot of people have looked at using generative AI in Excel to generate spreadsheet code. Among these applications are Data Sheets GPT, Excel Formula Bot (Formula Bot AI Excel formula generator), Google Workspace AI-Sheets, and Sheets AI. AI Office Bot not only generates formulas rapidly through text prompts, but it also explains them. This is particularly useful because Excel is one of the most commonly used tools in office work, and workflows could potentially see improvements in productivity.
Text-to-multilingual coding there are several software for multilingual code generation through just text inputs. Although ChatGPT is widely used for coding, there are many more generative AI applications that are being created for that purpose. While most of them work as coding assistants, they are also able to generate code through text prompts. Some of them are Alphacode (Li et al., 2023a-c), Amazon Codewhisperer, BlackBox AI, code complete, CodeGeeX, Codeium, GitHub Copilot, GitHub Copilot X, GhostWriter Replit, Mutable AI (mutable.ai. They are used to complete, explain, transform, and generate code. They generate new lines based on context and syntax. As we can observe, it is one of the fields with the biggest amount of applications. They can be personalized to your writing style; Cod
Applications such as AI2SQL and search AI can convert text into SQL code.
Word-to-website translation for topics like Mutiny, Durable, and website development. Both programs take text input and output a webpage with text and photos. In particular, we have three apps—Diagram AI, Galileo AI, and Uizard AI—that use Generative AI to generate excellent user interfaces and optimize the consumer experience. Even more impressively, the.com automates the process of creating web pages, allowing businesses to build unique profiles for each client.
New technologies in the realm of computer programming have arisen, some of which are unrelated to generative AI code. Locofy is one example of a design-to-code technology that may be used to create code for online and mobile applications. In addition, text-to-automation technologies are available via Drafter AI, a platform that can automate complex analytical activities, and Lasso AI, which can automate any process using natural language. Efforts by Even Adept to enable text-based user interaction with all aspects of a user's computer have recently surfaced.

Public Speaking
Artificial speech models attempt to replicate human speech. The development of speeches has been made simpler by text-to-speech technology. The use of generative AI has made voice cloning a breeze with other speech-to-speech technologies. Podcasts, YouTube videos, and assisting mutes in communication are just a few of the

many potential future uses for this technology.
The most significant models are VALL-E and speech-to-speech ones, such as ACE-VC (Hussain et al., 2023). VALL-E is able to transpose written words into speech with natural intonation and emotion based on the surrounding text using a three-second audio recording of a human voice.
In the same vein as the previously mentioned models, Supertone AI offers voice editing.
text-to-speech systems for the purpose of creating spoken language,
With the use of generative AI, it is now simple to record one's voice in response to text. Coqui, Descript Overdub, EllevenLabs Listnr, Lovo AI, Resemble AI, Replica Studios, Voicemod, and Wellsaid are just a few of the many platforms that have been developed. The most crucial model is AudioLM, which is Google's framework for consistently producing high-quality audio (Borsos et al., 2023) and is widely used. The Video Recordings-to-voice Dubverse is an excellent tool for video dubbing projects as it converts video recordings into voice.

Deep Learning for AI
The ability of AI to convert many forms of media into human-readable text, video, audio, and more has advanced to a satisfactory degree. This serves a significant purpose due to

machine learning's capacity to understand human speech and simplify it into simpler text. We can build more robust AI models if we can convert any input into text, which will make it easier to grasp and even use as input in other technologies. As an example, AI achieves an average inaccuracy of 1.15 percent in speech-to-text tasks, which is quite near to human performance (Radford et al., 2023). Computers will soon be able to allow rapid comprehension of huge and complicated datasets if they can transform all inputs into text, which is why these discoveries are quite encouraging. This area has enormous potential for the future of report generation, since it may include audio, graphics, and text into a written output that consumers will find very useful.
Modeling of generating regions into text has also made use of this technique. An object understanding transformer called GriT (Wu et al., 2022) uses pairings of regions and texts to find and characterize things. Object detection jobs may benefit from this.
There are other technologies that can convert photos to text as well. A number of domains may benefit from these technologies, including computer vision, which can aid AI in its comprehension of user-generated material. Alayrac et al. (2022), Kirillov et al. (2023), and VisualGPT (Chen et al., 2022) are a few examples of applications that include these technologies.
Various inputs-to-text Some systems even allow us to convert various inputs into text. Primer AI can process and respond to large volumes of media in real time, including text, photos, audio, and video. This information may be better understood and used to safeguard democracy and national security. With the use of speech AI, research and marketing teams may transcribe and analyze unstructured audio, video, and text to get competitive insights. Generative AI's ability to swiftly evaluate large and unstructured data sets is shown by both technologies. Primer can help us comprehend and act on it, and Speak AI can provide us insights rapidly.
Natural language is another tool that pdf-to-text generative AI may utilize to extract data from documents. The two programs are MapDeduce and ChatDOC. Using natural language searches, they may swiftly find what they need in PDFs and summarize it. This functionality has also been included into ChatGPT's model from late October 2023.
One of the most prominent areas in speech-to-text has been speech-
transcribing and subtitle technology are really beneficial. The following applications are included: Cogram AI, Deepgram AI, Dialpad AI, Fathom Video, Fireflies AI, GoogleUSM (Zhang et al., 2023a-e), Papercup, Reduct Video, Whisper (Radford et al., 2023), and Zoom IQ. It's important to note that not all of these technologies only execute speech-to-text jobs; in fact, some of them do many more. Using Deepgram AI, we can determine who is speaking,

syntax, as well as the search terms. The artificial intelligence features of Dialpad automate client touchpoints, provide real-time suggestions, and summarize calls. In addition to translating, Papercup can also generate voices that seem natural. The last point is that zoom has made use of AI in its systems, which includes tools like chat summaries and email composition. Workflow optimization may be seen by merging several generative AI technologies.
Another area where table data-to-text generative AI has proven effective is in turning data tables into text. Defog AI, MURMUR (Saha et al., 2022), and TabT5 are a few examples of Generative AI apps that serve this goal. Unstructured data may be understood by Murmur in particular. A significant improvement in the efficiency of company decision-making based on the rapid interpretation of table data is possible if we succeed in perfecting this technique.
This is something that videos-to-text flamingo can even handle when given video inputs. Both TwelveLabs and Minotaur were identified as video inputs (Goyal et al., 2023). TwelveLabs takes a video input and turns it into vector representations of important elements including objects, actions, voice, screen text, and people. You can do a fast search using these vectors. In its long-form films, Minotaur takes on query-based video comprehension. As it simulates the proper identification of visual scenes in movies, another model dubbed MOVIECLIP (Bose et al., 2023) was discovered to be particularly beneficial in this field. This innovation allows us to see the beginnings of

computers' ability to comprehend large, unstructured datasets.

Business

There is no doubt that generative AI will have far-reaching effects on the corporate world, making use of several of the aforementioned technologies in the process. Automating more creative and expensive operations, including creating marketing materials, PowerPoint presentations, or decreasing repetitive chores, may help firms save money. It has the potential to give rise to whole new kinds of AI-powered companies, like Truewind, which automates accounting, or Harvey, which automates litigation. We can only speculate as to the extent to which generative AI may impact business operations in the ways outlined below, given its relative youth. Goldman Sachs estimates that generative AI has the potential to propel GDP growth by 7%, highlighting the significance and influence of this technology.

Other sectors have also felt the effects, according to accounting records. Accounting is one area that has adopted Truewind, an AI-powered bookkeeper that aims to improve accuracy and transparency.

Generative AI may also be used to produce advertisements, as shown by several applications as Omneky, Pencil, Waymark, Clickable, and Ad Creative AI. Waymark, the final one, is great because it searches the web for local content and then uses that data to create films.

commercial information. Additionally, LensAI is helpful since it allows for more precise ad targeting by recognizing and pairing relevant advertising with specific items, logos, activities, and context. Generative AI may also support the storytelling in these commercials with apps like Subtxt and AI 21 Labs.

Another field that has made use of generative AI is architecture, namely with tools like Autodesk Space creator and SWAPP AI. SWAPP generates comprehensive, precise, and correct architectural construction documentation and building information modeling (BIM) files by using smart, sophisticated algorithms. Architects, city planners, and property developers may create first-rate site plans with the help of Autodesk Spacemaker, an AI program hosted in the cloud. Actually, Zuma—AI-powered real estate assistance that automates lead generation—shows that generative AI may also aid in the real estate aspect of the process.

From a research perspective, business visualizations are currently available in apps that allow you to ask basic questions using natural language and get responses supported by actual data. These responses are presented in a visual format that facilitates process integration and expedites market search. The following are examples of such businesses: Alphawatch, Dataherald, Maya, and OpenAxis AI.

In order to harness the potential of generative AI, the organization creates visually appealing content. As shown with Microsoft Designer, which generates invites, digital postcards, images, and more, designs may be swiftly made with just text prompts. One example of generative AI in action is Brandmark, which generates logos and other business-related material like business cards. Another example is Looka AI. Namelix, brandinition, and brandsnap are three resources that might help you generate ideas for company names.

Another use of customer assistance generative AI is the automation of customer communication. One Reach AI, OpenSight AI, Brainfish, and Yuma AI are some of the applications that may help your organization create customized chatbots. Generative AI also has tools like InboxPro, Lavender, Smartwriter, and Twain that can automate e-mails. Email analytics and social media data are two examples of the kinds of technology that might help improve business processes. There are also voice-assistant systems like Poly AI.

When it comes down to it, Generative

AI via Sonantic, an editing and directing tool for text-to-voice acting.

Additionally, educational institutions have been impacted. This is mostly due to the introduction of ChatGPT into students' assignments (Li et al., 2023a-c) and the rise of firms like Broadn, which use language models and generative AI to enable individuals to design personalized learning courses that align with their individual learning styles.

With the help of feedback-generative AI tools like Venturus AI and Mixo AI, we may get instantaneous feedback on our business ideas.

When it comes to the tedious process in the finance industry, generative AI can be a huge help. BloombergGPT is a finance-focused large-scale language model developed from the ground up. Sentiment analysis, news categorization, named entity identification, and question answering are just a few of its many potential applications. We can see how this may be very useful for anyone working in the financial sector. When it comes to modeling, Quilt Labs AI is a platform that uses AI to turn financial data into models. One sector that could benefit greatly from the automation that generative AI might provide is the financial sector.

Apps for internal communications, such as Albus ChatGPT, ChatGPT for Slack, and Moveworks, which automate employee assistance and provide discussion summaries, may greatly aid in staff management.

An excellent illustration of an industry that has been profoundly impacted by Generative AI is the legal documentation sector. Some examples of apps and businesses in this space are Spellbook Legal, Harvey AI, Casetext CoCounsel, and Darrow AI. Contract analysis, due diligence, litigation, and regulatory compliance are all areas where Harvey AI may be of assistance by providing data-driven insights, suggestions, and forecasts. Darrow

AI, on the other hand, may source cases for your legal business and do due diligence on them. In regards to Spellbook Legal, it evaluates and proposes the contract conditions using GPT-4. It is clear that Generative AI is taking over several tasks in the legal industry that might be mechanized with ease. An software called TaxGPT uses OpenAI's GPT to automate the process of filling out tax forms in this area.

Since generative AI may simplify the process of creating innovative regions and images, it has had a significant impact on marketing jobs. Anyword, Copy AI, Google Workspace-Gmail and Docs, Hyperwrite, Jasper, Letterdrop, Regie AI, Simplified AI, Type AI, and Writesonic are just a few of the many copywriting programs that have previously been created. Emails, website material, drafts, answers, marketing content, and product descriptions are all within the scope of capabilities. Many companies might benefit much from streamlining these procedures, as is readily apparent. To go the extra mile for your firm, Regie AI may modify the LLM's tone to match your brand's voice. Jasper can compose reports, emails, blogs, and social media postings; this is just one more example of how companies may enhance their operations by combining several generative AI technologies.

Companies like LA LA LAND, which uses an AI-powered digital model studio to display your 3D concepts as realistic models, are demonstrating how Generative AI is already impacting the modeling industry.

Through Salesforce's EinsteinGPT platform, which generates tailored content for all Salesforce clouds, we can see how firms are spending money on artificial intelligence (AI). All interactions with sales, service, marketing, commerce, and IT will produce content, revolutionizing the customer experience. A comprehensive program that streamlines all processes is AI Intern IO. It provides AI for a wide range of company activities, including text, reports, code, marketing, HR papers, legal documents, translations, and documentation. Generative AI applications, like as Cohere AI, which provides LLM for text retrieval, generation, and classification, may help enhance product design and product production.

Automating routine processes is another area where generative AI may be useful to businesses. Bardeen AI, Magical AI, and Notion AI are just a few of the numerous programs that can do this. programs like these are tailored to repetitive operations and are particularly helpful for businesses who are looking to automate relatively straightforward procedures using machine learning.

Thanks to the abundance of pre-existing applications, generative AI can boost sales. Applications that optimize contact centers and improve the customer experience include Cresta, Forethought AI, Grain AI, and Replicant. When it comes to customer care, Replicant is up to the task. Cresta and Grain are among the others that provide live support to call centers. Automatic note-taking, record-keeping, and insight-capture for client talks are automated by Grain AI, while Cresta turns real-time insights into real-time actions. When it comes to Forethought, their goal is to streamline the customer service process. In order to ensure that every sales call is well-prepared, the app Tennr was developed to provide ideal meeting prep materials. Copy Monkey AI is a software that can optimize your product's listing on Amazon and boost its organic ranking.

Similarly, it may facilitate laborious procedures used in scientific research. Applications like Koala, Elicit AI, ConsensusNLP, Agolo AI, and ArxivGPT demonstrate this. Finding articles, extracting important statements, and highlighting insights are some of the features of these programs. In particular, two chatbots tailored to the needs of scientists are Koala and ConsensusNLP. The use of generative AI for fact-checking has also been investigated by Golden.

Generative AI may help simplify the analyst's procedures by generating slides. Assisting with both slide creation and market research allows for this to be accomplished. Several applications exist that can generate presentations using natural language for the purpose of slide creation. Automatic Slides, Canva Docs to Decks, ChatBA, Decktopus, Gamma, Google Docs to Slides, Tome, and Slide AI are just a few of them.

Some of them, like Tome AI, just need a little textual instruction, while others, like Canva Docs, which turns papers into slide presentations, require lengthy texts to function. In addition, Decktopus can make valuable slide notes for you.

Apps like Clips AI, Pictory AI, Predis AI, Tweethunter, and Tweetmonk are designed to help in the development of social media content. Using AI, Clips AI and Pictory AI can transform lengthy pieces of material into shareable social media posts. Using your brand's voice, Predis AI creates visual and video content. Using either Tweethunter or Tweetmonk, you may get your brand's content tweeted. We can watch as Generative AI learns your brand's ins and outs and automates these tasks in no time. Through Bytepods, businesses may also take use of Generative AI to create their own podcasts.

Generative AI may also be useful for higher-level, more strategic divisions of a firm when making strategic decisions. Using GPT, applications such as Rationale AI may facilitate the development of many business assessments.

One last application of generative AI is to generate synthetic data that mimics real-world conditions for testing purposes. Hazy, Mostly AI, Octopize, and Tonic are some of the websites that may help you do this.

Additional features for trip planning include the ability to create itineraries; some examples of such applications are Roam Around, TripNotes, and the Kayak plug-in for ChatGPT. The capacity to build a visiting itinerary is shown by the first two, while the ability to search for

hotels, flights, and more using natural language is demonstrated by the Kayak plugin.

Video game

The ability to create generative AI from text, 3D models, and photos will be a huge boon to the game industry. Text models are useful for narrative and character development, whereas 3D models aid in the creative process. We can learn a lot about the potential of Generative AI in other industries by looking at the gaming industry as a case study. Since research indicates that the gaming industry would experience more realistic gaming environments and more engaging and individualized video games as a result of using Generative AI, these developments seem to be quite noteworthy (Wu et al., 2023). Your patience is much appreciated. The rearranged data is shown here.

When making video games, generative AI may be a useful tool. This is shown by means of programs such as Latitude, Illiad AI, and CSM. In particular, PixelVibe facilitates the production of game assets by use of generative AI.

Characters from video games: We've discovered Character AI, ConvAI, in-world AI, and RCT AI Chaos Box. Using natural language, ConvAI and World AI create characters. You can get entire characters only by adding character settings.

In terms of RCT AI Chaos Box, it is an engine that utilizes Deep Reinforcement Learning to assess player inputs in real-time and produce NPC replies and new narratives on the fly.

One game that uses LLMs for open-ended text-to-level creation is Levels MarioGPT (Sudhakaran et al., 2024).

Armorlab is a piece of software that uses artificial intelligence to fuel the creation of textures for video games.

Music

Generative AI may also significantly enhance the process of creating music. Inputting a simple text prompt or music tracks as input are both viable options for accomplishing this aim. Data from some experiments suggests that people can't tell the difference between music created by humans and music created by artificial intelligence. A song produced by Drake and The Weeknd using artificial intelligence had over 600,000 listens on Spotify in only one day, proving this point. Due to the Intellectual Property concerns that came from exploiting someone else's music samples for your own profit, it only lasted 24 hours. Here we see firsthand how issues related to the usage of publicly available data might emerge from irresponsible use of generative AI. There is a vast diversity of application cases for this technology that may enhance musicians' creativity beyond just making tunes. These tools may help artists overcome creative blocks by providing them with written suggestions, audio samples, and gestures to help spark ideas.

There have been several uses for text-to-music, which involves generating music using natural language instructions. The following systems are capable of producing music using basic natural language: Aiva, ERNIE-music (Zhu et al., 2023), Harmonai, Infinite Album, Jukebox, Mubert, Musico, Noise2Music (Huang et al., 2023), Sonify, soundful, and Splash AI. Based on language prompts, Splash AI may explore and mix samples, or it can use a sophisticated generative model trained on this dataset to generate original, high-quality music. Hyperpop, electronic dance music (EDM), glitch, phonk, trap, lo-fi, and hip-hop are just a few of the genres that users may make music in. The outcomes are remarkable even if the songs are just fifty seconds long. While the outcomes may vary, all of the models are comparable to Splash AI in that they generate musical compositions in response to a user-supplied natural language query.

Another kind of inputs-to-music model incorporates several kinds of

their technological advancements. In order to generate musical scores, musico responds to a variety of inputs, including motions, codes, and other noises. A new model called EDGE is making it possible to turn even dancing into music.

Applications like Moises AI and SingSong, which use Generative AI, may likewise enable musical editing (Donahue et al., 2023). The AI from Moises is useful.

practitioners of musical editing by removing vocals and altering tempo and timbre.

The field of biotechnology

Generative artificial intelligence tools facilitate the process of molecular modeling, which is useful in biotechnology. The fields of drug discovery and protein modeling may both benefit from this. Improvements in biotechnology may also be possible as a result of these technologies' continued development. The NASDAQ-listed Absci Corporation already employs generative AI in the development of new pharmaceuticals. Generative AI can rapidly generate realistic structures, increasing success rates to 20% (Winnifrith et al., 2023), which is a significant improvement over the relatively low success rates in protein modeling (Syrlybaeva and Strauch, 2023) for example. The potential for these technologies to significantly accelerate processes is clear.

Using NVIDIA Bionemo, a cloud service for generative AI in drug discovery, researchers have access to large-scale biomolecular AI models that can generate and forecast new compounds. Atomic AI, BigHat AI, Exscientia, Menten AI, and ProteinQure are just a few of the many firms that use Generative AI for medication innovation.

They develop antibody treatments for patients by using synthetic biology and machine learning, with BigHat AI serving as an example of their skills in the domain of drug

discovery. Through continual improvement, AI models enhance molecules.
Hundreds of distinct compounds may be designed simultaneously using their own machine-learning methods. The next step is to measure the molecular characteristics and add them to the model. Pharmaceutical firms achieve this level of efficiency via the use of Generative AI in conjunction with testing feedback. In addition to producing high-quality data, they both boost efficiency. It established a partnership with Merck to work together on up to three drug development projects in November 2022.
As for protein modeling, several models that have been discovered include BARTSmiles, a generative language model for molecular representation (Chilingaryan et al., 2022), and Alphafold, a computer program that predicts protein structures for the full human genome (Ren et al., 2023). Protein design using Generative AI is at the heart of two newly discovered firms, Cradle and Profluent.
To see an example of AI in action in Protein
concept, we may examine Cradle's commercial activities. To forecast the structure of a single sequence or a multimer, Cradle employs Alphafold's model. Utilizing models based on extensive datasets, Cradle enhances the thermostability of your proteins. Generative AI allows you to construct your own proteins, much as BigHat AI.

foretell how things will go and see the outcomes graphically. Once you have finished, you may share your findings with the firm to help them refine their models.

Brain
A brain model is a piece of software that can read and interpret brain waves. They have the potential to help us decipher the significance of a person's brain waves. Helping the deaf and hard of hearing communicate with generative AI is one of their most exciting potential applications. Even though it's still in its early stages of development, the field is seeing encouraging outcomes; for example, Meta's models achieved 41% accuracy. These findings hold great promise for altering the way doctors see brains. In addition, generative AI has the potential to improve classroom instruction. Technological advancements have been shown to facilitate education in recent studies (Darmawan, 2012). Researchers will have a better understanding of brain communication and be able to implement more efficient forms of communication if these technologies enable them to translate brain signals into text.
Using a Diffusion Model (DM) known as stable diffusion, a novel approach termed brain signals-to-images Using Stable Diffusion for Brain Images (Tang et al., 2023b) may convert brain waves into pictures. The participants' previously observed photographs might be recreated using this model. While the findings are excellent, it would take a lot of resources to retrain the model and do the experiment again. The use of steady diffusion to reconstruct brain pictures has never been done before.
When looking for models that can convert brain impulses into text, we came across Meta AI's brain speech and Non-Invasive Brain Recordings (Défossez et al., 2023). Both of them are attempting to decipher speech using non-invasive brain recordings.

Different kinds of
The models may integrate many technologies from the list and make the most of them. All of the apps on this list accept various inputs, which may speed up the development of AI. The future of Generative AI may lie in initiatives involving multi-tasking agents like GATO. While some models, like text-to-slides, do make use of several generative AI technologies, they were chosen for this reason alone.
A multimodal chatbot is one that can take in text, images, and audio as inputs and then output text and images. GPT-4 is the latest iteration of this technology. The sector is seeing a major shift as the most proficient chatbot transitions to multimodal. Chatbots' capabilities and the types of data they may assist users with are both enhanced when they accept several inputs at once. Increasing input types

have the potential to increase the technology's independence by allowing for the integration of AI into more operations than previously possible.
Kosmos-1, described by Huang et al. (2024), is one multimodal language model that has several capabilities. The ability to comprehend and generate new languages is one sort of capability. Perception-language activities include tasks like visual question answering, multimodal discourse, picture captioning, and image identification with descriptions. Prismer, as described in Liu et al. (2023a), is a paradigm for vision language that makes use of multi-modal experts. Question answering, object identification, segmentation, and picture captioning are a few examples of tasks. While using training data that is up to two orders of magnitude less, our model is comparable with state-of-the-art vision models that are currently available. The PaLM-E model is an embodied multimodal language model (Driess et al., 2023). First, PaLM-E's primary goal in development was to serve as a model for robotics; it accomplishes this goal by solving a wide range of tasks using a wide range of robot kinds and modalities, including pictures, robot states, and neural scene representations. Simultaneously, PaLM-E is an adequately competent vision-and-language model. It excels in both visual and verbal tasks, including object detection, scene classification, and quoting poetry, as well as solving mathematical problems and generating code. In the world of multi-format input chatbots, the ERNIE bot stands out. It can take a text prompt and turn it into text, graphics, audio, and video. What's more, it can even offer speech in many local dialects of Chinese, including Sichuan. Initial

expectations were not met by Ernie's user turnout. Still, the CEO has asserted that the chatbot can hold its own against ChatGPT.

The NVIDIA Picasso cloud service allows users to create and launch 3D, video, and picture apps that are driven by generative AI. It is part of a larger framework of integrated AI models for multimodal cloud services in generative AI. It combines text with images, video, and 3D models.

In fact, a framework known as HuggingGPT exists.

(Shen et al., 2024) that uses LLMs (see: ChatGPT) to link different AI models in ML communities (see: embracing face) in order to do AI tasks. With the use of language as a general interface, LLMs take on the role of controllers, managing existing AI models to complete complex AI tasks. Its remarkable performance in language, vision, and voice, among other difficult tasks, opens the door to more sophisticated AI. Using the same framework to combine many models may provide remarkable outcomes. Images with writing on them may be used to create sounds. The potential for AI to simultaneously optimize several industrial processes is almost limitless when these models are combined. An improvement in graphic design's efficiency might be achieved rapidly by simply merging text and picture models.

When it comes to Adobe FireFly, it's a family of models that takes text as input and outputs various media types such as photos, vectors, videos, and 3D models. It uses language commands to trigger the generation of pictures and other materials using many AI models. With Photoshop's new generative fill tool, users can easily add, expand, or delete material from photos using simple text prompts. Quickly removing people, adding puddles, and animals are just a few examples. It is also possible to make several drawings out of only one. Faster processes are the result of integrating all these functionalities into a single software.

When it comes to efforts to create generalist agents, GATO stands alone as an agent that goes beyond only text outputs. A generalist strategy that employs several modes, tasks, and embodiments is followed. Everyone on the same network can do things like push buttons, communicate, and play games simultaneously. Concerning Imbue (formerly generally intelligent), it is a business responsible for creating agents that are generally competent. They want to imbue machines with intellect and human values; they're worth $1 billion. Their goal is to implement AI systems that are matched with human levels, so that they can aid in scientific research and generalize to a variety of economically beneficial jobs. They do this by creating internal agent prototypes utilizing their large pre-trained models that are now focused on coding chores. This is the only firm that we are aware of that deals with AI agents. We also think their effort is important since Imbue is trying to make models that are easier to understand.

Artificial general intelligence (AGI) advancements are crucial because they bring computers closer to mimicking human intellect (Latif et al., 2023). By responding independently to human commands, AutoGPT became the first method for artificial general intelligence (Talebirad and Nadiri, 2023) in the field. As a result, the model's star rating on GitHub soared to over 150,000.

References

Alayrac, J. B., Donahue, J., Luc, P., Miech, A., Barr, I., Hasson, Y., ... & Simonyan, K. (2022). Flamingo: A visual language model for few-shot learning. *Advances in Neural Information Processing Systems*, *35*, 23716-23736. https://doi.org/10.48550/arXiv.2204.14198

Aydin, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. *Academic Platform Journal of Engineering and Smart Systems*, *11*(3), 118-134. https://doi.org/10.21541/apjess.1293702

Bahrini, A., Khamoshifar, M., Abbasimehr, H., Riggs, R. J., Esmaeili, M., Majdabadkohne, R. M., & Pasehvar, M. (2023, April). ChatGPT: Applications, opportunities and threats. *In 2023 Systems and Information Engineering Design Symposium* (*SIEDS*), (pp. 274-279). IEEE. https://doi.org/10.1109/SIEDS58326.2023.10137850

Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., ... & Fung, P. (2023). A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination and interactivity. *ArXiv Preprint ArXiv:2302.04023*. https://doi.org/10.48550/arXiv.2302.04023

Borsos, Z., Marinier, R., Vincent, D., Kharitonov, E., Pietquin, O., Sharifi, M., ... & Zeghidour, N. (2023). Audiolm: a language modeling approach to audio generation. *IEEE/ACM Transactions on Audio, Speech and Language Processing*. https://doi.org/10.1109/TASLP.2023.3288409

Bose, D., Hebbar, R., Somandepalli, K., Zhang, H., Cui, Y., Cole-McLaughlin, K., ... & Narayanan, S. (2023). Movieclip: Visual scene recognition in movies. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, (pp. 2082-2091). https://doi.org/10.1109/WACV56688.2023.00212

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of AI-generated content (aigc): A history of generative AI from gan to chatGPT. *ArXiv Preprint ArXiv:2303.04226*. https://doi.org/10.48550/arXiv.2303.04226

Chan, E. R., Nagano, K., Chan, M. A., Bergman, A. W., Park, J. J., Levy, A., ... & Wetzstein, G. (2023). Generative novel view synthesis with 3D-aware diffusion models. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, (pp. 4217-4229).

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. D. O., Kaplan, J., ... & Zaremba, W. (2021). Evaluating Large Language Models Trained on Code. *arXiv preprint arXiv:2107.03374*. https://doi.org/10.48550/arXiv.2107.03374

Chen, Z., Wang, G., & Liu, Z. (2023). Scenedreamer: Unbounded 3D Scene Generation from 2D Image Collections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2023.3321857

Cheng, C. Y., Huang, F., Li, G., & Li, Y. (2023). Play: Parametrically conditioned layout generation using latent diffusion. *arXiv preprint arXiv:2301.11529*. https://doi.org/10.48550/arXiv.2301.11529

Chilingaryan, G., Tamoyan, H., Tevosyan, A., Babayan, N., Khondkaryan, L., Hambardzumyan, K., ... & Aghajanyan, A. (2022). Bartsmiles: Generative masked language models for molecular representations. *arXiv preprint arXiv:2211.16349*. https://doi.org/10.48550/arXiv.2211.16349

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative Adversarial Networks: An overview. *IEEE Signal Processing Magazine*, *35*(1), 53-65. https://doi.org/10.1109/MSP.2017.2765202

Darmawan, D. (2012). Biological Communication Behavior through Information Technology Implementation in Learning Accelerated. https://doi.org/10.4236/ijcns.2012.58056

Donahue, C., Caillon, A., Roberts, A., Manilow, E., Esling, P., Agostinelli, A., ... & Engel, J. (2023). SingSong: Generating Musical Accompaniments from Singing. *arXiv preprint arXiv:2301.12662*. https://doi.org/10.48550/arXiv.2301.12662

Driess, D., Xia, F., Sajjadi, M. S., Lynch, C., Chowdhery, A., Ichter, B., ... & Florence, P. (2023). Palm-e: An Embodied Multimodal Language Model. *arXiv preprint arXiv:2303.03378*. https://doi.org/10.48550/arXiv.2303.03378

Défossez, A., Caucheteux, C., Rapin, J., Kabeli, O., & King, J. R. (2023). Decoding Speech Perception from non-invasive Brain Recordings. *Nature Machine Intelligence*, *5*(10), 1097-1107. https://doi.org/10.1038/s42256-023-00714-5

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). Gpts are gpts: An Early Look at the Labor Market Impact Potential of Large Language Models. *arXiv preprint arXiv:2303.10130*. https://doi.org/10.48550/arXiv.2303.10130

Fawzi, A., Balog, M., Huang, A., Hubert, T., Romera-Paredes, B., Barekatain, M., ... & Kohli, P. (2022). Discovering Faster Matrix Multiplication Algorithms with Reinforcement Learning. *Nature*, *610*(7930), 47-53. https://doi.org/10.1038/s41586-022-05172-4

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business and Information Systems Engineering*, *66*(1), 111-126. https://doi.org/10.1007/s12599-023-00834-7

Fridman, R., Abecasis, A., Kasten, Y., & Dekel, T. (2024). Scenescape: Text-Driven Consistent Scene Generation. *Advances in Neural Information Processing Systems*, *36*. pp. 39897-39914.

Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P., & Berner, J. (2024). Mathematical capabilities of chatgpt. *Advances in Neural Information Processing Systems*, *36*. pp. 27699-27744.

Gao, J., Shen, T., Wang, Z., Chen, W., Yin, K., Li, D., ... & Fidler, S. (2022). Get 3D: A generative model of high quality 3D textured shapes learned from images. *Advances in Neural Information Processing Systems*, *35*, 31841-31854.

Garrido-Merchán, E. C., Arroyo-Barrigüete, J. L., & Gozalo-Brihuela, R. (2023). Simulating HP Lovecraft Horror Literature with the ChatGPT Large Language Model. *arXiv preprint arXiv:2305.03429*. https://doi.org/10.48550/arXiv.2305.03429

Goyal, R., Mavroudi, E., Yang, X., Sukhbaatar, S., Sigal, L., Feiszli, M., ... & Tran, D. (2023). Minotaur: Multi-task Video Grounding from Multimodal Queries. *arXiv preprint arXiv:2302.08063*. https://doi.org/10.48550/arXiv.2302.08063

Hong, F., Chen, Z., Lan, Y., Pan, L., & Liu, Z. (2022). Eva3D: Compositional 3D Human Generation from 2D Image Collections. *arXiv preprint ArXiv: 2210.04888*. https://doi.org/10.48550/arXiv.2210.04888

Huang, Q., Park, D. S., Wang, T., Denk, T. I., Ly, A., Chen, N., ... & Han, W. (2023). Noise2music: Text-Conditioned Music Generation with Diffusion Models. *arXiv preprint arXiv:2302.03917*. https://doi.org/10.48550/arXiv.2302.03917

Huang, S., Dong, L., Wang, W., Hao, Y., Singhal, S., Ma, S., ... & Wei, F. (2024). Language is not all you need: Aligning Perception with Language Models. *Advances in Neural Information Processing Systems*, *36*. pp. 72096-72109.

Hussain, S., Neekhara, P., Huang, J., Li, J., & Ginsburg, B. (2023, June). Ace-vc: Adaptive and Controllable Voice Conversion using Explicitly Disentangled self-Supervised Speech Representations. *In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*), (pp. 1-5). IEEE. https://doi.org/10.1109/ICASSP49357.2023.10094850

Joublin, F., Ceravola, A., Deigmoeller, J., Gienger, M., Franzius, M., & Eggert, J. (2023). A Glimpse in ChatGPT Capabilities and its Impact for AI Research. *arXiv preprint arXiv:2305.06087*. https://doi.org/10.48550/arXiv.2305.06087

Kauers, M., & Moosbauer, J. (2022). The FBHHRBNRSSSHK-Algorithm for Multiplication in $\mathbb{Z}_2^{5\times5}$ is Still not the End of the story. *arXiv preprint arXiv:2210.04045*. https://doi.org/10.48550/arXiv.2210.04045

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment Anything. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, (pp. 4015-4026). https://doi.org/10.48550/arXiv.2304.02643

Koh, J. Y., Agrawal, H., Batra, D., Tucker, R., Waters, A., Lee, H., ... & Anderson, P. (2023). Simple and Effective Synthesis of Indoor 3D Scenes. *In Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. *37*, No. 1, pp. 1169-1178). https://doi.org/10.1609/aaai.v37i1.25199

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the News that's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation. *Journal of Experimental Political Science*, *9*(1), 104-117. https://doi.org/10.1017/XPS.2020.37

Kulkarni, C., Druga, S., Chang, M., Fiannaca, A., Cai, C., & Terry, M. (2023). A Word is Worth a Thousand Pictures: Prompts as AI Design Material. *arXiv preprint arXiv:2303.12647*. https://doi.org/10.48550/arXiv.2303.12647

La Salvia, M., Torti, E., Leon, R., Fabelo, H., Ortega, S., Martinez-Vega, B., ... & Leporati, F. (2022). Deep Convolutional Generative Adversarial Networks to Enhance Artificial Intelligence in Healthcare: A skin cancer application. *Sensors*, *22*(16), 6145. https://doi.org/10.3390/s22166145

Latif, E., Mai, G., Nyaaba, M., Wu, X., Liu, N., Lu, G., ... & Zhai, X. (2023). Artificial General Intelligence (AGI) for Education. *arXiv preprint arXiv: 2304.12479*, *1*. https://doi.org/10.48550/arXiv.2304.12479

Leiker, D., Gyllen, A. R., Eldesouky, I., & Cukurova, M. (2023). Generative AI for Learning: Investigating the Potential of Synthetic Learning Videos. *arXiv preprint arXiv2304.03784*. https://doi.org/10.48550/arXiv.2304.03784

Lewkowycz, A. andreassen, A., Dohan, D., Dyer, E., Michalewski, H., Ramasesh, V., ... & Misra, V. (2022). Solving Quantitative Reasoning Problems with Language Models. *Advances in Neural Information Processing Systems*, *35*, 3843-3857.

Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023a). ChatGPT in Education: A Discourse Analysis of Worries and Concerns on Social Media. *Education and Information Technologies*, 1-34. https://doi.org/10.1007/s10639-023-12256-9

Li, Y., Liu, H., Wu, Q., Mu, F., Yang, J., Gao, J., ... & Lee, Y. J. (2023b). Gligen: Open-set grounded text-to-image generation. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 22511-22521). https://doi.org/10.48550/arXiv.2301.07093

Li, C., Zhang, C., Waghwase, A., Lee, L. H., Rameau, F., Yang, Y., ... & Hong, C. S. (2023c). Generative AI Meets 3D: A Survey on Text-to-3D in AIGC Era. *arXiv preprint arXiv:2305.06131*. https://doi.org/10.48550/arXiv.2305.06131

Lin, C. H., Gao, J., Tang, L., Takikawa, T., Zeng, X., Huang, X., ... & Lin, T. Y. (2023). Magic 3D: High-resolution text-to-3D content creation. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 300-309). https://doi.org/10.48550/arXiv.2211.10440

Liu, J., Snodgrass, S., Khalifa, A., Risi, S., Yannakakis, G. N., & Togelius, J. (2021). Deep Learning for Procedural Content Generation. *Neural Computing and Applications*, *33*(1), 19-37. https://doi.org/10.1007/s00521-020-05383-8

Liu, S., Fan, L., Johns, E., Yu, Z., Xiao, C., & Anandkumar, A. (2023a). Prismer: A Vision-Language Model with an Ensemble of Experts. *arXiv preprint arXiv:2303.02506*, *3*. https://doi.org/10.48550/arXiv.2303.02506

Liu, G. H., Vahdat, A., Huang, D. A., Theodorou, E. A., Nie, W., & Anandkumar, A. (2023b). I^2SB: Image-to-IMAGE Schr\" Odinger Bridge. *arXiv preprint arXiv:2302.05872*. https://doi.org/10.48550/arXiv.2302.05872

Melas-Kyriazi, L., Laina, I., Rupprecht, C., & Vedaldi, A. (2023). Realfusion: 360° reconstruction of any object from a single image. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, (pp. 8446-8455).

Murphy, K. P. (2022). *Probabilistic Machine Learning: An Introduction*. MIT press. ISBN-10: 0262046822.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187-192. https://doi.org/10.1126/science.adh2586

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of AI on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590*. https://doi.org/10.48550/arXiv.2302.06590

Pepe, M., Alfio, V. S., & Costantino, D. (2023). Assessment of 3D Model for Photogrammetric Purposes Using AI Tools Based on NeRF Algorithm. *Heritage*, *6*(8), 5719-5731. https://doi.org/10.3390/heritage6080301

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). Robust Speech Recognition Via Large-Scale Weak Supervision. *In International Conference on Machine Learning*, (pp. 28492-28518). PMLR. https://proceedings.mlr.press/v202/radford23a.html

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, *1*(2), 3. https://3dvar.com/Ramesh2022Hierarchical.pdf

Ren, F., Ding, X., Zheng, M., Korzinkin, M., Cai, X., Zhu, W., ... & Zhavoronkov, A. (2023). AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. *Chemical Science*, *14*(6), 1443-1452. https://doi.org/10.1039/d2sc05709c

Ren, X., & Wang, X. (2022). Look outside the room: Synthesizing a consistent long-term 3D scene video from a single image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 3563-3573). https://doi.org/10.48550/arXiv.2203.09457

Saha, S., Yu, X. V., Bansal, M., Pasunuru, R., & Celikyilmaz, A. (2022). Murmur: Modular multi-step reasoning for semi-structured data-to-text generation. *arXiv preprint arXiv:2212.08607*. https://doi.org/10.48550/arXiv.2212.08607

Schick, T., Dwivedi-Yu, J., Jiang, Z., Petroni, F., Lewis, P., Izacard, G., ... & Riedel, S. (2022). Peer: A collaborative language model. *arXiv preprint arXiv:2208.11663*. https://doi.org/10.48550/arXiv.2208.11663

Schwitzgebel, E., Schwitzgebel, D., & Strasser, A. (2023). Creating a large language model of a philosopher *Mind and Language*. https://doi.org/10.1111/mila.12466

Shen, Y., Song, K., Tan, X., Li, D., Lu, W., & Zhuang, Y. (2024). Hugginggpt: Solving AI tasks with chatgpt and its friends in hugging face. *Advances in Neural Information Processing Systems*, *36*, pp. 38154-38180.

Singer, U., Sheynin, S., Polyak, A., Ashual, O., Makarov, I., Kokkinos, F., ... & Taigman, Y. (2023). Text-to-4D dynamic scene generation. *arXiv preprint arXiv:2301.11280*. https://doi.org/10.48550/arXiv.2301.11280

Singhal, K., Tu, T., Gottweis, J., Sayres, R., Wulczyn, E., Hou, L., ... & Natarajan, V. (2023). Towards expert-level medical question answering with large language models. *arXiv preprint arXiv:2305.09617*. https://doi.org/10.48550/arXiv.2305.09617

Sison, A. J. G., Daza, M. T., Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023). ChatGPT: More than a "weapon of mass deception" ethical challenges and responses from the Human-Centered Artificial Intelligence (HCAI) perspective. *International Journal of Human–Computer Interaction*, 1-20. https://doi.org/10.1080/10447318.2023.2225931

Sudhakaran, S., González-Duque, M., Freiberger, M., Glanois, C., Najarro, E., & Risi, S. (2024). Mariogpt: Open-ended text2level generation through large language models. *Advances in Neural Information Processing Systems*, *36*. pp. 54213-54227.

Syrlybaeva, R., & Strauch, E. M. (2023). Deep learning of protein sequence design of protein–protein interactions. *Bioinformatics*, *39*(1), btac733. https://doi.org/10.1093/bioinformatics/btac733

Talebirad, Y., & Nadiri, A. (2023). Multi-agent collaboration: Harnessing the power of intelligent llm agents. *arXiv preprint arXiv:2306.03314*. https://doi.org/10.48550/arXiv.2306.03314

Tang, J., LeBel, A., Jain, S., & Huth, A. G. (2023a). Semantic Reconstruction of Continuous Language from Non-Invasive Brain Recordings. *Nature Neuroscience*, *26*(5), 858-866. https://doi.org/10.1038/s41593-023-01304-9

Tang, J., Wang, T., Zhang, B., Zhang, T., Yi, R., Ma, L., & Chen, D. (2023b). Make-it-3D: High-fidelity 3D creation from a single image with diffusion prior. *In Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 22819-22829). https://make-it-3d.github.io/

Taylor, R., Kardas, M., Cucurull, G., Scialom, T., Hartshorn, A., Saravia, E., ... & Stojnic, R. (2022). Galactica: A large language model for science. *arXiv preprint arXiv:2211.09085*. https://doi.org/10.48550/arXiv.2211.09085

Tevet, G., Raab, S., Gordon, B., Shafir, Y., Cohen-Or, D., & Bermano, A. H. (2022). Human motion diffusion model. *arXiv preprint arXiv:2209.14916*. https://doi.org/10.48550/arXiv.2209.14916

Tian, H., Lu, W., Li, T. O., Tang, X., Cheung, S. C., Klein, J., & Bissyandé, T. F. (2023). Is ChatGPT the ultimate programming assistant-how far is it?. *arXiv preprint arXiv:2304.11938*. https://doi.org/10.48550/arXiv.2304.11938

Weng, C. Y., Srinivasan, P. P., Curless, B., & Kemelmacher-Shlizerman, I. (2023). Personnerf: Personalized reconstruction from photo collections. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 524-533). https://doi.org/10.48550/arXiv.2302.08504

Winnifrith, A., Outeiral, C., & Hie, B. (2023). Generative artificial intelligence for de novo protein design. *arXiv preprint arXiv:2310.09685*. https://doi.org/10.48550/arXiv.2310.09685

Wu, J., Wang, J., Yang, Z., Gan, Z., Liu, Z., Yuan, J., & Wang, L. (2022). Grit: A generative region-to-text transformer for object understanding. *arXiv preprint arXiv:2212.00280*. https://doi.org/10.48550/arXiv.2212.00280

Wu, Y., Yi, A., Ma, C., & Chen, L. (2023). Artificial intelligence for video game visualization, advancements, benefits and challenges. *Mathematical Biosciences and Engineering*, *20*(8), 15345-15373. https://doi.org/10.3934/mbe.2023686

Xu, D., Jiang, Y., Wang, P., Fan, Z., Wang, Y., & Wang, Z. (2023). Neurallift-360: Lifting an in-the-wild 2d photo to a 3D object with 360° views. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 4479-4489).

Yu, J., Xu, Y., Koh, J. Y., Luong, T., Baid, G., Wang, Z., ... & Wu, Y. (2022). Scaling autoregressive models for content-rich text-to-image generation. *arXiv preprint arXiv:2206.10789*, *2*(3), 5. https://3dvar.com/Yu2022Scaling.pdf

Vahdat, A., Williams, F., Gojcic, Z., Litany, O., Fidler, S., & Kreis, K. (2022). Lion: Latent point diffusion models for 3D shape generation. *Advances in Neural Information Processing Systems*, *35*, pp. 10021-10039.

Zhang, M., Qamar, M., Kang, T., Jung, Y., Zhang, C., Bae, S. H., & Zhang, C. (2023a). A survey on graph diffusion models: Generative AI in science for molecule, protein and material. *arXiv preprint arXiv:2304.01565*. https://doi.org/10.48550/arXiv.2304.01565

Zhang, Y., Han, W., Qin, J., Wang, Y., Bapna, A., Chen, Z., ... & Wu, Y. (2023b). Google USM: Scaling automatic speech recognition beyond 100 languages. *arXiv preprint arXiv:2303.01037*. https://doi.org/10.48550/arXiv.2303.01037

Zhang, C., Zhang, C., Zhang, M., & Kweon, I. S. (2023c). Text-to-image diffusion model in generative Ai: A survey. *arXiv preprint arXiv:2303.07909*. https://doi.org/10.48550/arXiv.2303.07909

Zhang, C., Zhang, C., Zheng, S., Qiao, Y., Li, C., Zhang, M., ... & Hong, C. S. (2023D). A complete survey on generative AI (AIGC): Is chatGPT from GPT-4 to GPT-5 all you need? *arXiv preprint arXiv: 2303.11717*. https://doi.org/10.48550/arXiv.2303.11717

Zhang, C., Zhang, C., Zheng, S., Zhang, M., Qamar, M., Bae, S. H., & Kweon, I. S. (2023e). Audio diffusion model for speech synthesis: A survey on text to speech and speech enhancement in generative AI. *arXiv preprint arXiv:2303.13336*. https://doi.org/10.48550/arXiv.2303.13336

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*. https://doi.org/10.48550/arXiv.2303.18223

Zhu, P., Pang, C., Chai, Y., Li, L., Wang, S., Sun, Y., ... & Wu, H. (2023). ERNIE-music: Text-to-waveform music generation with diffusion models. *arXiv preprint arXiv:2302.04456*. https://doi.org/10.48550/arXiv.2302.04456