



≡ Menu

## W1: Introduction & Capabilities of Generative AI

# Lesson Summary

Congratulations! You have completed this lesson.

At this point, you have learned about the fundamentals and evolution of Generative AI. You have identified the capabilities of Generative AI and explored their applications in the real world. You even gained some hands-on experience using Generative AI use cases.

Specifically, you learned that:

- Generative AI models can generate new content based on the data they are trained on.
- While discriminative AI mimics our analytical and predictive skills, Generative AI also mimics our creative skills.
- The creative skills of Generative AI are built on the use of models, such as GANs, VAEs, and transformers.
- Foundation models in Generative AI can be adapted to create specialized models or tools tailored to specific use cases.
- Generative AI models and tools have a broad scope for applications across different domains and industries.
- Generative AI can create coherent and contextually relevant content and generate realistic, high-quality images, synthetic voices, new audio, and dynamic videos.
- Generative AI models can generate and complete code and synthesize new data to augment existing datasets.
- Generative AI models can create highly realistic and complex virtual worlds, including avatars and digital personalities.

[Go to next item](#)

✓ Completed

# Lesson Summary

## W2: Applications & Tools for GenAI.

Congratulations! You have completed this lesson.

At this point, you have learned about some of the applications of Generative AI across domains. You explored some common Generative AI tools for text, image, and code generation. You even had the opportunity to see the generation in action through hands-on lab experiences.

Specifically, you learned that:

- Generative AI has been applied to numerous fields, including IT and DevOps, entertainment, finance, medicine, and human resources.
- In the workplace, the application of Generative AI is changing how we work and making us more efficient and successful in what we do.
- Generative AI has the capability for text generation. Text generation tools can generate different kinds of text, translate languages, and answer questions.
- There are two popular **text generation tools**: OpenAI's **ChatGPT** and **Google Gemini**. ChatGPT is more effective for maintaining conversational flow, and Gemini is a better choice for research.
- Generative AI tools can also produce **new images** through text or image prompts. These image generation tools can transform images through techniques like **image-to-image translation, style transfer, inpainting, or outpainting**.
- Some common image generation models include **DALL-E**, **Stable Diffusion**, and **StyleGAN**. In addition, numerous Generative AI tools are available for image generation and transformation.
- Microsoft provides **Bing Image Creator** for image generation, and **Adobe Firefly** is a family of Generative AI tools for Adobe's Creative Cloud applications.
- Generative AI tools can generate new code, produce optimized versions of existing code, or convert code from one programming language to another.
- A few prominent code generators are **GPT**, **GitHub Copilot**, **PolyCoder**, and **IBM Watson**. ChatGPT and Gemini can be valuable tools for basic code generation. However, these tools have limitations based on their training.

**GitHub Copilot, PolyCoder & IBM Watson : Prominent Code generators**

- |   |   |
|---|---|
| <ul style="list-style-type: none"><li>• <b>Real-time generators</b></li><li>• <b>Code restructuring</b></li><li>• <b>Solution templates</b></li></ul> | <ul style="list-style-type: none"><li>• Accelerate development cycle</li><li>• Promoting coding best practices</li><li>• Foster consistent coding standards</li></ul> |
|---|---|

# History and Evolution of Generative AI

## Objectives

After completing this reading, you will be able to:

- Describe the key developments in the field of generative AI in chronological order
- Explain the current scenario of generative AI

## Introduction

Generative artificial intelligence (AI) is a field that focuses on developing algorithms to create new content, such as text, images, music, and code. Generative AI models are trained on substantial datasets of existing content and learn to generate new content similar to the data they were trained on.

## History

The origin of generative AI origins can be traced back to the initial stages of artificial intelligence exploration. In the 1950s, researchers began to explore the use of computers to generate new data, such as text, images, and music. However, the computational power and data resources needed for these systems to flourish were not yet available.

One of the earliest instances of generative AI dates back to 1964 with the creation of the **ELIZA chatbot**. Operating on a rule-based system, ELIZA simulated conversations with users by generating responses based on received text. While not genuinely intelligent, ELIZA showcased the potential of generative AI for human-like communication.

During the 1980s and 1990s, hardware and software capabilities advanced considerably and facilitated the development of advanced generative AI models, including **neural networks**. Neural networks are inspired by the human brain and can learn intricate patterns in data. However, these early neural networks were computationally expensive to train and could only generate small amounts of content.

In the early 2000s, a significant breakthrough occurred in generative AI research with the advent of **deep learning**. Utilizing neural networks with multiple layers, deep learning models could be trained

on extensive datasets to discern complex patterns, enabling the generation of new data that closely resembled human-created content. This breakthrough led to the development of innovative generative AI models, including generative adversarial networks (GANs) and variational autoencoders (VAEs).

GANs and VAEs excel at producing high-quality content that is often indistinguishable from content crafted by humans.

**GANs** operate by training two neural networks in opposition: a generator that creates new content and a discriminator that tries to differentiate between real and synthetic content. Eventually, the generator learns to craft content realistic enough to deceive the discriminator.

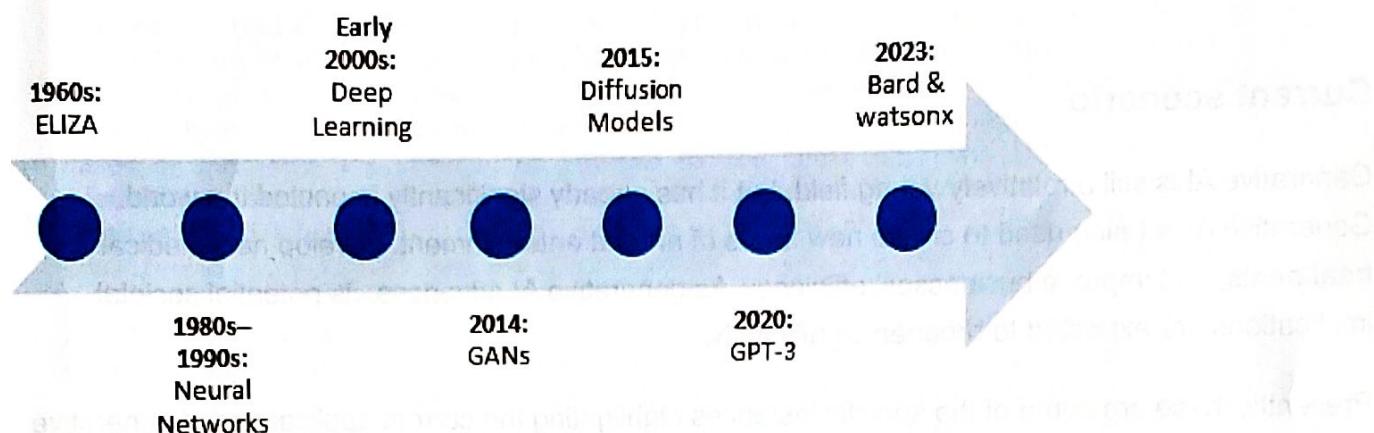
**VAEs** work by learning a latent space of the data they are trained on. The latent space is a representation of the data that captures the most essential features of the data. VAEs can generate new content by sampling from the latent space and decoding the latent code into the original data space.

In recent years, there has been a rapid explosion in the development of new generative AI models. These models can now generate a wide variety of content, including text, images, music, and code. Generative AI is also used in various applications, such as art, design, and healthcare.

One such instance is the development of **diffusion models** in 2015. Diffusion models work by gradually adding noise to a clean image until it is completely unrecognizable. They can then be reversed to gradually remove the noise and generate a new image. Diffusion models have been used to create high-quality images and text.

The next significant development is that of **large language models (LLMs)** like GPT-3 and Gemini in 2020 and 2023, respectively. LLMs are trained on massive datasets of text and code, which allows them to generate realistic text, translate languages, write different kinds of creative content, and answer your questions in an informative way. 2023 was also when **watsonx**, a superior generative AI platform based on the cloud, was introduced by IBM. Watsonx can support multiple LLMs.

Here is a brief **timeline** of some of these milestones in the fascinating journey of generative AI:



## 1960s: ELIZA

ELIZA, an early chatbot, showcased early attempts at simulating conversation.

## 1980s–1990s: Neural network development

Researchers started developing more sophisticated generative AI models, including neural networks, capitalizing on advances in hardware and software.

## Early 2000s: Deep learning

Deep learning, a breakthrough in AI, gained prominence. Neural networks with multiple layers were employed for training on massive datasets.

## 2014: Generative adversarial networks (GANs)

Introduced by Ian Goodfellow and his colleagues, GANs presented a revolutionary two-player neural network framework.

## 2015: Diffusion models

The development of diffusion models brought a novel approach to image generation by gradually adding noise to a clean image.

## 2020: GPT-3

OpenAI released GPT-3, a state-of-the-art language model, demonstrating impressive natural language understanding and generation capabilities.

## 2023: Gemini and Watsonx

Another large language model, Google's Gemini, and IBM's generative AI system, WatsonX are introduced, further advancing the capabilities of generative AI.

These milestones represent an overview of the generative AI journey, capturing vital developments in natural language processing, image generation, and the underlying architectures that have shaped the field over the years.

## Current scenario

Generative AI is still a relatively young field, but it has already significantly impacted the world. Generative AI is being used to create new forms of art and entertainment, develop new medical treatments, and improve businesses' efficiency. As generative AI advances, its potential societal implications are expected to broaden significantly.

Presently, these are some of the specific instances highlighting the current applications of generative AI:

- **Art and entertainment:** Generative AI is being used to create new art forms, such as AI-generated paintings, music, and literature. Generative AI is also being used to develop new video games and other interactive experiences.
- **Medicine:** Generative AI is being used to develop new medical treatments, such as personalized cancer therapies and AI-powered drug discovery. Generative AI is also being used to develop new medical imaging tools and improve diagnosis and treatment accuracy.
- **Business:** Generative AI is being used to improve the efficiency of businesses by automating tasks such as customer service, marketing, and sales. Generative AI is also being used to develop new products and services.

Generative AI holds significant transformative potential across various facets of our lives. Using generative AI responsibly and ethically is essential, but it is also important to be excited about its possibilities.

## Summary

In this reading, you learned about the history of generative AI. You noted the key developments in chronological order. You learned about the contribution of each key development to the field of generative AI. Finally, you learned about the current scenario of generative AI.

## W2a : Text Generation

Welcome to Tools for Text Generation. After watching this video, you'll be able to describe the basics of text generation through generative AI, explain the key capabilities of common models and tools for text generation. At the core of the text generation capabilities of generative AI are Large Language Models or LLMs. Based on patterns and structures learned during training, LLMs interpret context, grammar and semantics to generate coherent and contextually appropriate text. Drawing statistical relationships between words and phrases allows LLMs to adapt creative writing styles for any given context. LLMs are the basis for many text generation models. Two such examples are Generative Pre-trained Transformer or GPT and Gemini AI model. These models have evolved into multi-modal models, offering multiple capabilities. Let's learn about the capabilities of these models through two popular tools, ChatGPT and Google Gemini. ChatGPT is based on GBT as the Large Language Model and uses Advanced Natural Language Processing or NLP. While originally, ChatGPT only took text prompts as input to generate new contents, with the newer versions, it can take both image and text inputs. ChatGPT offers diverse capabilities for text generation. It is capable of smooth and context-based conversations. Let's start a conversation with ChatGPT to learn a concept. Input a prompt that says, I've heard about generative AI and want to learn more. ChatGPT responds with some basic information based on the context. When you take the conversation forward to refine the research by asking, how can I use generative AI to improve my storytelling skills? As a prompt, ChatGPT provides the response based on the context and question provided by you. Feel free to experiment and guide the conversation further. ChatGPT will build an informative and interesting conversational flow. It can also help you with varied creative tasks. Let's enter the prompt, help me to create slides to demonstrate the features of a learning platform and ChatGPT comes with suggestions about the title, content, and visuals for specific slides. Although ChatGPT is most proficient in English, it can understand and respond to several other languages. Let's prompt it to write hello in French and Spanish, and it generates the desired output. ChatGPT can also be a useful tool to assist you in learning a new language or any subject for that matter. Another popular text generation tool is Google Gemini. It is powered by Google's Gemini AI model, which is a new family of multimodal AI models developed by Google DeepMind and designed for advanced reasoning, understanding, and content generation across text, images, audio, video, and code. Built for high efficiency and scalability, Gemini is optimized for multimodal learning, enabling seamless interaction across different formats. It is designed to handle multiple types of data and tasks within a single model, making it a powerful tool for diverse AI applications. Try experimenting with different prompts to explore the capabilities of Google Gemini. Let's try Google Gemini, the prompt to get a summary of the latest news on a topic, such as provide a summary of the latest news on the war in Ukraine. It provides you with multiple drafts as the response. You can select one of these or regenerate. Next, let's try Google Gemini, to generate ideas or solve a problem. It's prompted to provide a strategy for a digital marketing campaign for promoting a fashion brand. It provides a step-by-step approach to the marketing campaign. ChatGPT and Google Gemini offer capabilities for other valuable use cases. For example, they can help you with basic mathematics, statistics, and problem solving through these subjects. They are also proficient in financial analysis, investment research, budgeting, and more. Furthermore, ChatGPT and Google Gemini can generate code and perform code-related tasks across various programming languages and frameworks. Having interacted with both ChatGPT and Google Gemini,

you'll notice that ChatGPT is more effective in generating dynamic responses and maintaining conversational flow, while Google Gemini may be a better choice for researching the latest news or information on a topic as it has access to web sources through Google Search and Google Scholar. It's important to realize that generative AI models, including GPT and Gemini AI model, are evolving, so their capabilities and features may change. Apart from ChatGPT and Google Gemini, there are other text generators as well. Jasper, for example, generates high-quality marketing content of any length, tailored to a brand's voice. Writer is a valuable tool for creating high-quality content for blogs, emails, SEO metadata, and ads on social media. Also, Copy.ai is great for creating content for social media, marketing, and product descriptions. Another tool, WriteSonic, offers specific templates for different types of text, such as articles in blogs, ads, and marketing. There are also tools available for specific use cases. For example, tools like Resumer can generate a summary of a text by extracting key ideas or concepts. Next, tools like UClassify are used for classification to assign one or more categories to a snippet of text. Tools for sentiment analysis understand and generate text that reflects the underlying emotions expressed in human language. Examples include Brand24 and Repustate. For multilingual language translation, you can use LanguageWeaver and Yandex. It is important to note that many of the open-source generative AI tools collect and review the data shared with them to improve their systems. This is an important consideration for interacting with these tools to avoid sharing any confidential or sensitive information. So, do we have open-source privacy-preserving alternatives? The answer is yes. GPT4ALL, for example, can be installed on your machines to run as a privacy-aware chatbot without internet or a graphics processing unit. Further, chatbots like H2OAI and PrivateGPT are designed to protect user privacy by running on local machines, without any internet connection using the power of LLMs. Not only that, you can customize these tools for use within a specific organization by linking them to your organization's documents and databases. Generative AI-based text generators offer several benefits. These tools are good learning aids as they provide step-by-step explanations. They can generate different forms of text quickly, enabling efficiency for writers and creators. These tools enhance creativity and inspire new ideas. By enabling engaging and interactive conversations, they are useful as virtual assistants and chatbots. By automating repetitive writing tasks, they can increase productivity for organizations. With multilingual support, they enable communication and content localization for global audiences. In this video, you learned that LLMs interpret context, grammar, and semantics to generate coherent and contextually appropriate text. LLMs are the basis for many text generation tools. Two popular text generation tools are OpenAI's ChatGPT and Google Gemini.

## Hands-on Lab

## W2b : Image Generation

Welcome to Tools for

Image Generation. After watching this video, you'll be able to describe the basic capabilities of

generative AI models for image generation and explain the key capabilities of common models and tools

for image generation. Generative AI image

generation models can generate new images and customize real and generated images to give you the desired output. For example, you may

want to generate an image of a child with

a book in her hand. Further, you may want

to change the color of the book cover in

the generated image. Let's generate a new image using a free AI image

generator, Freepik. You need to enter a text prompt describing the image you want to create. Let's say you enter

the following prompt, a boat sailing on a

calm lake at sunset, surrounded by lush

greenery and a serene sky. Remember, how you describe your image and the words you include in the prompt determine

the accuracy and quality of the image

that gets generated. Let's select the style

and generate the image. Here, we have multiple

images generated. You can select and

download an image, or you may want to generate other images by modifying the prompt. Let's look at some

more possibilities of image generation models. **Image to image translation** refers to transforming

an image from one domain to another while preserving the original matter and style. For example, converting

sketches to realistic images, converting satellite images to maps, converting security

camera images to higher resolution images, and enhancing detail and medical imaging. **Style transfer and fusion**

involve extracting the style from one image

and applying it to another, creating hybrid or fusion images, for example, converting a

painting to a photograph. **Inpainting** refers to reconstructing missing or damaged parts of

an image to make it complete. You can use this for art restoration, forensics, removing unwanted

objects and images while preserving continuity and context, and blending virtual objects into real world scenes and

augmented reality. **Outpainting** involves extending

the original image by generating new parts to it that are like extensions of the original. This can be used for

generating larger images, enhancing resolution, and

creating panoramic views. The image generation and

modifications capabilities of generative models and tools have evolved with the evolution

of models that power them. OpenAI's **DALL-E** is

based on the GPT model. Trained on larger datasets of images and their textual descriptions, DALL-E can generate high resolution images in multiple styles, including photorealistic images and paintings. DALL-E has evolved in the new versions of DALL-E to provide capabilities for generating multiple image variations and image transformation through inpainting and outpainting. Stable Diffusion is an open source text to image diffusion model. Diffusion models are generative models that can create high resolution images. Stable Diffusion is primarily used to generate images based on text prompts. Though it can also be used for image to image translation, inpainting and outpainting. NVIDIA StyleGAN model separates the modeling of image content and image style, enabling precise control over style for manipulating specific features like pose or facial expression. StyleGAN has evolved to generate higher resolution images with more realistic details. You can explore generative AI's text to image generation capabilities using free tools like Craiyon, Freepik and Picsart. These tools can generate images in different forms and styles. Fotor and Deep Art Effects offer a variety of pretrained styles allowing you to create your own custom styles. DeepArt.io is an online platform that turns photos into artwork of different styles. Midjourney is a platform that enables image generator communities that help artists and designers to create images using AI and explore each other's creations. Many generative AI image generators can also be integrated as APIs to embed their functionality and capabilities into different software programs and tools. Some popular image generators that offer APIs include DALL-E, Midjourney, and Craiyon. Technology giants such as Microsoft and Adobe have also stepped into the world of AI image generators. Microsoft Bing Image Creator is based on the DALL-E model. You can access this tool by navigating to bing.com/create or through Microsoft Edge. This makes Microsoft Edge the first browser with an integrated AI image generator. Adobe Firefly is a family of generative AI tools designed to integrate with Adobe's Creative Cloud applications, such as Photoshop and Illustrator. Firefly is trained on Adobe stock photos, openly licensed content, and public domain content. Firefly can take text prompts in over 100 languages and include tools that allow you to manipulate color, tone, lighting, composition, generative fill, text effects, generative recolor, 3D to image, and extend image. In this video, you learned that generative AI based models and tools can generate new images through both text

## W2c: Tools for Audio and Video Generation

Welcome to Tools for audio and video generation. After watching this video, you'll be able to describe how generative AI audio and video tools create impactful media content, explain the key capabilities of generative AI audio and video tools, explore generative AI's ability to re-imagine virtual worlds. Market.us estimates that the generative AI music market, valued at 229 million in 2022, will register a high CAGR of 28.6% to reach 660 million by 2032. Generative AI music is created using generative AI audio capabilities. Over the past few years, these capabilities are helping companies and individuals novice or experienced, simplify their processes to bring their complicated visions to life. Think about this. Suppose you've been putting off starting your podcast or adding some sound effects to your remixes. In that case, you'll love what generative AI audio tools can do for you. They come in three categories, speech generation tools, music creation tools, and tools that enhance audio quality. **Speech generation tools** are mostly text to speech or TTS tools that convert text into audio. While read aloud technology is not new, generative AI architecture has upgraded how this technology works. Deep learning algorithms are repeatedly trained on vast datasets of human speech. This allows them to break down and efficiently replicate vocal characteristics, such as pronunciation, speed, emotion, and intonation. As a result, generative AI TTS tools create more accurate natural sounding speech, which is especially helpful to those who struggle with visual impairment, language barriers, and other reading disabilities. On the fun side, these tools can help you listen to essays, feedback, and notes, which might be easier than reading them. They can also help you communicate better. What if you wish to narrate your presentation in a standout manner? You could log into LOVO, Synthesia, Murf.ai, or Listenr, and choose from vast libraries of AI voices, languages, or emotions. You could even **create a unique voice or clone your voice**. Some tools will also let you edit your vocal tracks pronunciation, tone, and speed to create a professionally sounding final product. **What about music?** Let's say one sunny afternoon, the amateur musician in you is feeling motivated. You could try Meta's AudioCraft, the generative AI tool pretrained on sound effects in 20,000 hours of Meta-owned or licensed music. There's also Shutterstock's Amper Music, AIVA, Soundful, Google's Magenta, and the GPT-4-powered WavTool. These tools let you choose from extensive music banks, different music genres, instrumental styles,

and melodies. All you need to do is enter a text prompt. Based on your request, the tool will write short melodies or riffs, suggest or add instruments, compose a new song, or create a soundtrack for your next YouTube or Instagram video. Generative AI can also help you mix, master, and publish your final musical output on popular streaming platforms. You can even use audio enhancing tools. These are pretrained to identify specific sounds and can add fun sounds to your audio or remove unwanted ones. For example, Descript can help you remove background noise, enhance low quality recordings, and add the desired sound effects. Auto AI cleans your files of unwanted noise. Many music generation tools also possess audio editing and enhancement capabilities. However, some projects need more than eclectic sound effects. In 2022, Runway AI used generative AI capabilities to produce the Oscar winning movie, Everything Everywhere all at once. Even if you're not making big cinema, you can use generative AI video tools in your everyday life. Let's say you're making a documentary on the lack of trees in your city. You could log into Runway's Gen-1 tool, which transforms existing video clips into different styles or use Runway's Gen-2 tool to create a video using text image or video inputs. Alternatively, you can use the EaseUS video toolkit or the Synthesia app. These tools will allow you to upload photos. If you don't have any, use text prompts to generate the images you need. Additionally, you can use these tools to record a narration, enhance your audio, convert your video file format, and publish your video. Synthesia even allows you to create custom avatars to increase your brand recall. Generative AI can enhance your virtual world experience. You can create unique imaginative virtual worlds with hybrid characteristics and exotic landscapes. Generative models can also respond in real time, improving the accuracy of simulations. Metaverse platforms employ generative AI to create a more personalized and engaging user experience. Gaming metaverses allow you to rapidly generate 3D objects and even create avatars fitted with specific personality traits that reflect in their expressions, behaviors, conversations, and decisions. The Sandbox, for example, is a metaverse where users can instantly build own, and market their games globally. Scenario AI helps create mobile gaming assets. In this video, you learned how generative AI audio and video tools can make an impact. With a simple text prompt, you can produce human-sounding speech in multiple languages, record songs, add sound effects, or

## W2d: Tools for Code Generation

Welcome to Tools for Code Generation. After watching this video, you'll be able to describe the basic capabilities of generative AI for code generation. Discuss the strengths and limitations of text-generating tools for code generation. And explain the key capabilities of common models and tools for code generation. Generative AI models and tools for code generation can generate code based on natural language input. Based on Deep Learning and Natural Language Processing, or NLP, these models comprehend the context and produce contextually appropriate code. Code generators can generate a new code snippet or a program from a text prompt. They can predict lines of code to complete a partial code snippet. They can also produce optimized versions of existing code. Further, these code generators can convert code from one programming language to another. They can also generate summaries and comments for code to improve documentation. Code generators can also recommend a complete programming solution. Describe a problem you're trying to solve, and code generators will suggest algorithms, data structures, and the right programming approaches. Let's explore the capabilities of GPT for code generation. Common AI's GPT excels in human-like text generation and demonstrates impressive capabilities in code creation. Here's an example of generating simple Python code through ChatGPT, a tool based on GPT. When you enter a text prompt, write a Python code to generate a message to greet a person. ChatGPT generates the Python code for it. Interestingly, it also provides guidelines about how to run this code. For generating code, it's important that you provide a clear prompt, specify the programming language, and provide other relevant requirements and constraints. To demonstrate how GPT helps debug code, enter an incorrect code as a text prompt in ChatGPT. An incorrect code and text are prompted to ChatGPT. ChatGPT provides the correct code and explanations of the corrections made. GPT also enables translating code from one programming language to another. It can also generate code documentation and comments to improve readability. GPT-based models and tools have evolved to generate longer and more accurate codes. This enables using these models and tools for coding to develop apps, websites, and plugins. Furthermore, the evolution of GPT makes it possible to generate code from images. For example, you can input the image of a course outline to generate code for a fully functional app. Google Gemini also offers code generation and debugging capabilities in more than 20 programming languages. ChatGPT and Gemini are valuable tools for coding or a new programming language because they provide a step-by-step and detailed explanation for better understanding. ChatGPT and Gemini excel at generating code with basic logic and programming concepts. They may not, however, be able to generate large or complex code from scratch. And while these tools understand programming concepts and syntax, they may not completely understand semantics. So, while the generated code is technically accurate, it may still not function as required. It's important to note that the knowledge of these models is limited to the data they're trained on. A specific version of GPT may not be aware of programming frameworks and libraries released after its training. For example, the knowledge of GPT 3.5 is limited to September 2021. So, if you want a more dated code generator, consider using models and tools designed specifically for code generation. GitHub Copilot, an AI code generator, can generate code based on various programming languages and frameworks. GitHub Copilot is powered by OpenAI Codex, which is a generative, pre-trained language model. It helps developers generate solution-based code. Copilot is trained on natural language text and source code from publicly available sources including GitHub repositories. Copilot can be integrated as an extension with popular code editors, such as Visual Studio.

Copilot can produce code snippets that adhere to best practices and industry standards. Polycoder is an open-source AI code generator. It's based on GPT-2 and trained on data from various GitHub repositories, written in 12 programming languages. It's particularly accurate for writing C codes. Polycoder offers an extensive library of pre-defined templates that can be used as blueprints for code generation for various use cases. Polycoder can help create, review, and refine code snippets precisely customized to requirements. Different code generators offer specific functionality and features. However, when the requirement is to enable hybrid cloud developers to write code for diverse needs, IBM Watson Code Assistant is an option. It's built on IBM Watson X.AI Foundation models for developers of any skill level. You can integrate Watson Code Assistant with a code editor. It enables developers to write code accurately and efficiently with real-time recommendations, autocomplete features, and code restructuring assistance. Moreover, you can input code or project files into Watson Code Assistant for analysis. It identifies patterns, suggests improvements, and generates code snippets or templates. Developers can customize this generated code for specific project needs. Many other AI-driven code generators and code assistant tools are available to help developers write accurate code faster. Amazon Code Whisperer can be integrated with code editors and offers real-time code recommendations. Tab9 can be helpful for accurate code completions. Repl.it is a platform that provides an interactive space for users to code, learn, and collaborate. With automatic code writing and optimization capabilities, AI-based code generators help developers increase their productivity and quality of code. They enable rapid prototyping to iterate on design ideas. These tools also help with cross-platform compatibility and migration by enabling multi-language code translation. AI-based code generators follow consistent patterns and coding standards. They can suggest refactoring patterns to adhere to best practices. However, these tools should be used with caution to ensure that the AI-generated code does not lead to ethical issues. For example, security vulnerabilities as these tools may be used for generating malicious data or data biases based on training. In this video, you learned that generative AI-based models and tools can generate new code from text and image prompts, optimize existing code, and produce solution-based code. ChatGPT and Gemini are useful for simple code generation, debugging, and learning programming. Prominent code generators like GitHub Copilot, Polycoder, and IBM Watson Code Assistant offer varied features such as real-time recommendations, code restructuring, and solution templates. In general, code generators improve productivity, accelerate development cycles, promote coding best practices, and foster consistent coding standards.

## Generative AI: Introduction and Applications

### Glossary

Welcome! This alphabetized glossary contains many of the terms in this course. This comprehensive glossary also includes additional terms not used in course videos. These terms are essential for you to recognize for better comprehension of the concepts covered in the course.

**Estimated reading time:** 3 minutes

Term	Definition
<b>Data augmentation</b>	A technique commonly used in machine learning and deep learning to increase the diversity and amount of training data.
<b>Deep learning</b>	A subset of machine learning that focuses on training computers to perform tasks by learning from data. It uses artificial neural networks.
<b>Diffusion model</b>	A type of generative model that is popularly used for generating high-quality samples and performing various tasks, including image synthesis. They are trained by gradually adding noise to an image and then learning to remove the noise. This process is called diffusion.
<b>Discriminative AI</b>	A type of artificial intelligence that distinguishes between different classes of data.
<b>Discriminative AI models</b>	Models that identify and classify based on patterns they observe in training data. In general, they are used in prediction and classification tasks.
<b>Foundation models</b>	AI models with broad capabilities that can be adapted to create more specialized models or tools for specific use cases.
<b>Generative adversarial network (GAN)</b>	A type of generative model that includes two neural networks: generator and discriminator. The generator is trained on vast data sets to create samples like text and images. The discriminator tries to distinguish whether the sample is real or fake.
<b>Generative AI</b>	A type of artificial intelligence that can create new content, such as text, images, audio, and video.
<b>Generative AI models</b>	Models that can understand the context of input content to generate new content. In general, they are used for automated content creation and interactive communication.
<b>Generative pre-trained transformer (GPT)</b>	A series of large language models developed by OpenAI. They are designed to understand language by leveraging a combination of two concepts: training and transformers.
<b>Large language models (LLMs)</b>	A type of deep learning model trained on massive amounts of text data to learn the patterns and structures of language. They can perform language-related tasks, including text generation, translation, summarization, sentiment analysis, and more.
<b>Machine learning</b>	A type of artificial intelligence that focuses on creating algorithms and models that enable computers to learn and make predictions or decisions. It involves designing systems that can learn from training data.
<b>Natural language processing (NLP)</b>	A branch of artificial intelligence that enables computers to understand, manipulate and generate human language (natural language).
<b>Neural networks</b>	Computational models inspired by the structure and functioning of the human brain. They are a fundamental component of deep learning and artificial intelligence.
<b>Prompt</b>	Instructions or questions that are given to a generative AI model to generate new content.
<b>Training data</b>	Data (generally, large datasets that also have examples) used to teach a machine learning model.
<b>Transformers</b>	A deep learning architecture that uses an encoder-decoder mechanism. Transformers can generate coherent and contextually relevant text.
<b>Variational autoencoder (VAE)</b>	A type of generative model that is basically a neural network model designed to learn the efficient representation of input data by encoding it into a smaller space and decoding back to the original space.



# Skills Network

# Hands-on Lab: Text Generation in Action (Lab 2a)

Estimated time needed: 30 minutes

Welcome to the hands-on lab text generation in action!

## What is text generation in generative AI?

Generative AI's greatest potential is to generate context-based and captivating content, leading to limitless possibilities for intelligent text generation.

Generative AI-powered models and tools can be used to generate a wide range of text formats, such as articles, blog posts, social media posts, product descriptions, marketing copy, language translation, conversational interactions, and image descriptions, among many more.

In this lab, you will look at the two most popular text generation tools, OpenAI's ChatGPT and Google's Gemini and explore how to use them for text generation and text summarization.

## Learning Objectives

After completing this lab, you will be able to:

- Explore the text generation capability of generative AI
- Write emails using ChatGPT
- Convert the unit measurement using ChatGPT
- Summarize text using Google's Gemini

## ChatGPT overview

ChatGPT is an AI chatbot of OpenAI based on GPT's large language models. Generative AI technology understands your natural language prompt and then uses the context to build the response and generate the output.

## Exercise 1: Write an email using OpenAI's ChatGPT

Writing an effective email for the required subject and context requires cognitive attention, as you must consider the affirmative language, response aligned with the trail mail, grammar, and so on.

### Step 1: Log in steps and the home page

1. Launch ChatGPT and sign in.

Note: If you cannot access ChatGPT by clicking on the link, copy the following URL and paste it into a browser.

<https://chat.openai.com/>

④<sub>1</sub>: You are a customer support rep at XYZ Inc. You are following up with a customer for feedback... Write a feedback request email...

2. Once launched, you must sign up or login.

Note: You can create a free account on ChatGPT and use GPT-3.5 features for free.

# Lab: Image Generation in Action (Lab 2b)



**Skills**  
Network

Estimated time: 30 minutes

Welcome to the hands-on lab image generation in action.

## What is image generation in generative AI?

Images are vital to visual details because they comprehensively represent concepts, data, and information that users can easily grasp. For effective communication, images are powerful elements that add depth, context, and visual appeal to the information.

Generative AI models have advanced significantly in recent years, enabling the creation of new and realistic images using artificial intelligence techniques. Using generative AI-powered models and tools, you can generate and modify high-quality images that are difficult to distinguish from real photographs.

This lab focuses on generative AI's image-generation capabilities. In this lab, you will look at how images are generated in Generative AI Classroom using the Dall-E 3 model. You will explore the platform and use it to create images from textual descriptions.

## Learning Objectives

After completing this lab, you will be able to:

- Explore the image-generation capabilities of generative AI
- Prompt generative AI models to generate the desired images
- Create realistic images from text in the Generative AI Classroom using the Dall-E 3 model
- Use different scenarios to generate images using the Generative AI Classroom

## About generative AI classroom lab

The generative AI classroom allows you to write and compare prompts to generate the desired text with real-time chat responses. Moreover, you can choose from multiple generative AI models and learn about their strengths and weaknesses.

You will have the generative AI classroom environment and instructions on one page in a browser. The instructions will be on the left half of the screen, and the generative AI classroom will be on the right half of the screen. You can interact with the language model using the message and chat fields.

The screenshot shows a split-screen interface. On the left, under the heading 'INSTRUCTIONS', there is a sidebar with icons for Home, Skills Network, Estimated time, What is image generation in generative AI?, Learning Objectives, and a Chat section. The main content area displays the 'Hands-on-Lab: Image Generation in Action' section, which includes the estimated time (30 minutes), a welcome message, a question about what image generation is, an answer about its importance, a description of the lab's goal, and a list of learning objectives. On the right, under the heading 'Generative AI Classroom', there is a sidebar with icons for New, Compare Models, Chat (selected), and Freeform. The main content area displays the 'How to use AI Classroom' section, which includes four steps: 1. Name the chat, 2. Choose your model(s), 3. Set the prompt instructions, and 4. Chat with your language model. There is also a note about elevating the AI learning experience by checking out IBM Watson AI now!

## Exercise 1: Create realistic images from text in the Generative AI Classroom

In this exercise, you will create realistic images using the Dall-E 3 model based on the textual prompts you provide to the generative AI tool. Dall-E 3 is a foundation model where you can input a short prompt, and the system generates a complex image based on your parameters. It takes English text as an input, called the "text prompt" and generates images based on the text description.

### Step 1: Set up the AI classroom

# Hands-on Lab: Image Generation in Action with Microsoft Copilot and Microsoft Designer(Optional) (Lab 2c)

Estimated time: 45 minutes

Welcome to the hands-on lab image generation in action with Microsoft Copilot and Microsoft Designer.

## What are AI image generation tools?

AI image generation tools are programs that use advanced technology to create or improve images without human intervention. These tools can create lifelike images, improve blurry pictures, remove unwanted elements, and create images based on written descriptions. Popular AI image generation tools like Microsoft Copilot, DALL-E, and StyleGAN are widely used in art, design, and research to make image-making process easier and create exceptional visuals.

This lab focuses on generative AI's image-generation capabilities. You will examine and explore the Microsoft Copilot/Microsoft Designer and use it to generate image captions and create thumbnails from textual descriptions.

## Learning Objectives

After completing this lab, you will be able to:

- Explore the image-generation capabilities of generative AI using Microsoft Copilot and Microsoft Designer
- Generate an image caption using Microsoft Copilot
- Create thumbnail images representing larger images using Microsoft Copilot
- Generate a social media post using Microsoft Designer

Note: These exercises are optional, as using the Microsoft Copilot/Designer requires signing in with your email account.

## Exercise 1: Generate image caption using Microsoft Copilot

In the digital age, image captions have gained even more importance due to their role in improving accessibility for people with visual impairments. Screen readers and other assistive technologies can read these captions aloud, allowing individuals who cannot see the image to understand its content. Generative AI models simplify the task of predicting the captions for the given image.

In this exercise, you will generate the caption for an image using GPT-powered Microsoft Copilot.

### What is Microsoft Copilot?

Microsoft Copilot is an LLM-based (large language model) chatbot developed by Microsoft. It can cite sources, create poems, and write songs.

Copilot uses generative AI technology (Microsoft Prometheus model) built upon OpenAI's GPT-4 foundational large language model, which has been fine-tuned using supervised and reinforcement learning techniques. Copilot's conversational interface style resembles ChatGPT's. The chatbot can communicate using numerous languages and dialects.

The free version of Microsoft Copilot includes the standard Copilot chatbot.

Microsoft Copilot is oriented toward generating images from your ideas that enhance your design or even create clever caption and hashtag options.

Let's generate captivating captions for the images using Microsoft Copilot's GPT model.

### Step 1: Select the image for caption generation

Select the image for which you want to generate the caption.

### Step 2: Launch Microsoft Copilot

1. Click [Microsoft Copilot](#) to launch it.

Note: If you cannot access Microsoft Copilot by clicking the link, copy the following URL and paste it into a browser.

<https://copilot.microsoft.com/>