**DALL-E** (stylized as **DALL·E** ) and its successors **DALL-E 2** and **DALL-E 3** are [computer programs](https://de.wikipedia.org/wiki/Computerprogramm) developed by [OpenAI](https://de.wikipedia.org/wiki/OpenAI) that can create images from text descriptions using [machine learning](https://de.wikipedia.org/wiki/Maschinelles_Lernen) . The name is a [portmanteau](https://de.wikipedia.org/wiki/Kofferwort) of the small animated robot Wall-E from the [film of the same name](https://de.wikipedia.org/wiki/WALL%C2%B7E_%E2%80%93_Der_Letzte_r%C3%A4umt_die_Erde_auf) and the Spanish [surrealist](https://de.wikipedia.org/wiki/Surrealismus)[Salvador Dalí](https://de.wikipedia.org/wiki/Salvador_Dal%C3%AD) . [[ 1 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-:0-1) The program uses [artificial neural networks](https://de.wikipedia.org/wiki/K%C3%BCnstliches_neuronales_Netz) to transfer words as input into arrays of [pixels](https://de.wikipedia.org/wiki/Pixel) as output. Using text arrays, the program can create [photorealistic images.](https://de.wikipedia.org/wiki/Fotorealismus)[[ 2 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-2) The [artificial intelligence](https://de.wikipedia.org/wiki/K%C3%BCnstliche_Intelligenz) can map completely new concepts and create images in various artistic styles. [[ 3 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-3)*In order to generate images, the model was trained using millions of images available on the Internet. The program is based on the Generative Pre-trained Transformer 3 (GPT-3)* , also developed by OpenAI , a text generator that can create texts, text summaries, and even poems. [[ 4 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-4)

**technology**

[ [Edit](https://de.wikipedia.org/w/index.php?title=DALL-E&veaction=edit&section=1) | [Edit source](https://de.wikipedia.org/w/index.php?title=DALL-E&action=edit&section=1) ]

Image generated by DALL-E 2.0 for the text “A [Shiba Inu](https://de.wikipedia.org/wiki/Shiba_Inu) dog wearing a beret and black turtleneck”

The Generative Pre-Trained Transformer (GPT) model was originally developed by [OpenAI](https://de.wikipedia.org/wiki/OpenAI) in 2018 using a Transformer architecture. [[ 5 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-KIPedia1-5) The first iteration, GPT, was scaled to produce GPT-2 in 2019; it was scaled again in 2020 to produce GPT-3 with 175 billion [parameters](https://de.wikipedia.org/wiki/Parameter_(K%C3%BCnstliche_Intelligenz)) . DALL-E's model is a multimodal [implementation](https://de.wikipedia.org/wiki/Implementierung) of GPT-3 with 12 billion parameters, "swapping text for pixels," trained with text-image pairs from the web. DALL-E 2 uses 3.5 billion parameters, fewer than its predecessor. [[ 5 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-KIPedia1-5)

*A conceptual representation of the Wikipedia encyclopedia* . Prompt: "Create an image of the Wikipedia encyclopedia according to your own ideas (but not a logo)."

DALL-E was developed and released in conjunction with CLIP (Contrastive Language-Image Pre-training). CLIP is a separate model based on zero-shot learning, trained on 400 million image pairs with text labels from the internet. Its task is to "understand and rank" the output of DALL-E by predicting which label is most appropriate for an image from a list of 32,768 labels randomly selected from the dataset (one of which was the correct answer). This model is used to filter a larger initial list of images generated by DALL-E to select the most appropriate outputs.

DALL-E 2 uses a diffusion model conditioned on CLIP image embeddings generated during inference from CLIP text embeddings by a prior model.

**Version overview**

[ [Edit](https://de.wikipedia.org/w/index.php?title=DALL-E&veaction=edit&section=2) | [Edit source](https://de.wikipedia.org/w/index.php?title=DALL-E&action=edit&section=2) ]

DALL-E was introduced by OpenAI on January 5, 2021. [[ 6 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-6) The announcement stated that the software was still in the research phase and access was limited to pre-selected [beta users](https://de.wikipedia.org/wiki/Betaversion) . The model can still make serious errors, even those that no human would make. [[ 7 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-7) The improved version, DALL-E 2.0, was introduced in April 2022. [[ 8 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-8) The less powerful version, *Craiyon* (initially known as *DALL-E mini* ), which is based on the DALL-E [source code](https://de.wikipedia.org/wiki/Quellcode) , has been publicly available since April 21, 2022. The model's surreal images were promptly used by internet users to create [memes](https://de.wikipedia.org/wiki/Meme_(Kulturph%C3%A4nomen)) . [[ 9 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-:1-9)[[ 10 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-10)

The general availability of DALL-E 2.0 was announced in September 2022. [[ 11 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-11)[[ 12 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-12)

DALL-E 3 was released in October 2023 for ChatGPT Plus and ChatGPT Enterprise customers. [[ 13 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-13)

**Opinions on DALL-E**

[ [Edit](https://de.wikipedia.org/w/index.php?title=DALL-E&veaction=edit&section=3) | [Edit source](https://de.wikipedia.org/w/index.php?title=DALL-E&action=edit&section=3) ]

The model's performance was described as "impressive" [[ 1 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-:0-1) and "uncanny" [[ 14 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-14) . Some observers saw it as a sign of the strong progress in [machine learning](https://de.wikipedia.org/wiki/Machine_Learning) . Developers of OpenAI and [Sam Altman](https://de.wikipedia.org/wiki/Sam_Altman) , the CEO, declared DALL-E to be a crucial step on the way to [Artificial General Intelligence](https://de.wikipedia.org/wiki/Artificial_General_Intelligence) . Others disagreed, saying that DALL-E did not have intelligence comparable to that of humans, as it did not demonstrate any general understanding of concepts. Furthermore, it would still make many errors due to a lack of understanding of objects in the physical world. [[ 15 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-15) The model could, however, have a strong influence on artists and graphic designers. [[ 16 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-16) The generation of disinformation and fake images was feared in an article in the [*New York Times*](https://de.wikipedia.org/wiki/New_York_Times) . [[ 17 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-17) The program could also have a cultural bias, as it is based solely on the [English language](https://de.wikipedia.org/wiki/Englische_Sprache) and the images used to train the program are largely from Western cultures, meaning that the generated images could therefore reflect [stereotypical ideas.](https://de.wikipedia.org/wiki/Stereotyp)[[ 9 ]](https://de.wikipedia.org/wiki/DALL-E#cite_note-:1-9)