

OpenAI Deep Research

了解です。「4o Image Generation」について、以下の観点で詳しく調べてまとめます：機能の概要、使い方、技術的な仕組み、他の画像生成サービスとの比較、リリース日とアップデート内容、従来の画像生成機能との違いなど。調査が完了したら、わかりやすくお伝えしますね。

ChatGPT の画像生成・編集機能「4o Image Generation」調査レポート

1. 機能の概要

「4o Image Generation」は、ChatGPT の GPT-4o モデルに統合された最新の画像生成・編集機能です。テキストによる指示から多彩な画像を生成できるほか、会話を通じて画像を逐次編集・改良することも可能です ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#)) ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。以下に主な特徴をまとめます。

- **高度な画像生成:** 風景・人物の写真風画像からイラスト・漫画風のスタイルまで、幅広い画像を生成できます ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。GPT-4o は従来の DALL-E 3 よりも強力な画像生成能力を備えており、**フォトリアルな描写も可能です** ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。実際、OpenAI は「市場で最高のモデルになり得る」と評しており、画像の美しさと実用性を両立しています ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。
- **テキストの正確な描画:** 生成した画像内に文字(看板の文、メモの手書き文字、ロゴなど)を**自然かつ正確に描くことができます** ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。例えば、日本語で「3月29日 休校」と書かれたメモ画像も鮮明に生成可能です ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#)) ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。従来

は日本語の文字が潰れたり意味不明になることが多かったのが、大幅に改善されています ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。

([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))生成された手書きメモ風の画像。「メモに手書きの日本語で『3月29日 休校』と書いてください」というプロンプトに対する出力例 ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))

- **複雑な指示への対応:** プロンプト(指示文)の解釈力が飛躍的に向上し、細かな構図指定や多数の要素を含むシーンも忠実に再現します ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。例えば「青い空の下、赤い自転車に乗った少年が手を振っている」と指示すれば、まさにその内容の画像を生成します ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。複数のオブジェクトや人物の関係性も正確に描写でき、OpenAI によれば最大 10~20 個の異なるオブジェクトを扱えるといいます ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#)) (従来モデルでは 5~8 個程度が限界でした)。
- **チャット形式での画像編集:** 生成後の画像に対して追加の指示を送り、段階的に修正・洗練することができます。例えば最初に自転車に乗る少年の画像を生成し、続けて「背景を雨にして」と指示すれば、同じ少年・スタイルのまま背景だけ雨天に変更できます ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。さらに「スタイルを変えずに、この少年が転んでしまった画像を生成して」と追記すれば、少年が転倒した新たな画像に更新されます ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。このように一貫性を保ちながら複数回にわたり画像を出力・編集でき、キャラクターデザインの微調整など試行錯誤がしやすくなっています ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。

[\(ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速\)](#)マルチターン編集の例。最初に「青空の下で自転車に乗る少年」の画像を生成し、続けて「スタイルを変えずに少年が転んでしまった画像を」と指示して得られた結果 ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#)) ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))

- **画像を入力に用いた生成:** ユーザーが画像をアップロードし、そのスタイルや内容を参照して新たな画像を生成することもできます ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。例えば料理の写真を与えて「この料理を半分食べかけにしてください」と指示すれば、元写真の文脈を踏まえた変換画像を得られます。GPT-4o モデルは画像とテキストの同時分布データで訓練されており、**画像同士の関係性**も学習しているため、アップロード画像の細部や文脈を解析してシームレスに統合することが可能です ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。
- **広範な知識の活用:** GPT-4o は ChatGPT (GPT-4) の持つ世界知識や推論能力をそのまま画像生成に活かします ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。例えばコードや数式を解釈してその内容を図として描いたり、科学実験の概略図、レシピの手順イメージなど**実用的で意味のある画像**を生成できます ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。単に写真風の美麗さだけでなく、説明図やインフォグラフィックなど業務に役立つ図も得意で、OpenAI も**業務関連の画像**(インフォグラフィックや図表)の生成に重点を置いたと説明しています ([「ChatGPT」、4o モデルによる画像生成が可能に - 無料プランでも - CNET Japan](#))。
- **多様なスタイル対応:** 写実写真風、アニメ・漫画風、水彩画風など、プロンプトで指定することで自由に**画像のテイスト**を変化させられます ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。訓練データに様々な画像スタイルが含まれているため、デザイン用途では求めるテイストに合わせた出力を得やすく、**創作にも実務にもそのまま使える品質**を実現しています ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。

以上のように、「4o Image Generation」により ChatGPT は文章と言葉だけでなく画像も生成・編集できるマルチモーダル AI へと進化しました。画像生成 AI としての表現力・実用性が飛躍的に向上し、初心者でも日本語で直感的にビジュアルコンテンツを作成できるようになっています。

2. 使い方

利用方法はシンプルで、ChatGPT のモデルとして「GPT-4o」を選択し、チャット欄に生成したい画像の内容を指示するだけです ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。例えば「〇〇の画像を作ってください」と日本語で入力すれば、GPT-4o がプロンプトを解釈して画像を生成します。日本語で直接指示できるため、複雑な英語プロンプトや専門知識は不要です ([ChatGPT の画像生成 AI 完全ガイド：無料でも使える最新機能を徹底解説](#))。実際、「帽子をかぶった猫の写真を生成して」といった簡単な依頼でも問題なく応えてくれます。

画像の生成結果は、通常数十秒～1 分以内でチャット画面に表示されます ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 — 画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。処理が完了すると「画像が作成されました」とともに画像が貼り付けられ、逐次読み込まれて完成します ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。表示された画像をクリックすると拡大表示でき、右上のダウンロードボタンから保存することも可能です ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。

逐次編集を行う場合は、同じスレッド内で追加のメッセージを送り、画像に対する変更点を指示します。例えば「もう少し明るくして」「背景を夕焼けに変更して」などと追加入力すれば、GPT-4o が直前の画像を踏まえて修正した新しい画像を返します ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。この際、スタイルや構図の一貫性は保たれるため、一からプロンプトを書き直す必要はありません ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。この会話的な編集操作により、理想の画像に近づくまで何度でも微調整を重ねることができます ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。

ユーザー自身の画像を編集することも可能です。ChatGPT の画像アップロード機能 (📎 アイコン) を用いて写真や画像ファイルをアップし、「この画像から〇〇を削除して」「この写真を油絵風にしてください」といった指示を与えると、GPT-4o が入力画像を解析して指定通りに変換・編集しま (ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch)】。アップロード画像を視覚的なインスピレーションとして活用し、新たな画像を生成することも得意で (ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch)】。このように、テキストによるインペインティング (画像内の編集) や画像変換をチャット形式で手軽に行える点も大きな魅力です。

ユーザーインターフェース (UI) は従来の ChatGPT と同じチャット画面で、テキストの入力欄にプロンプトを打つだけです。特別なソフトやコマンドは不要で、対話の流れで「…な画像を作って」とお願いすれば ChatGPT が自動で画像モードに切り替わりま (ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速)】。モデル名が GPT-4o になっていることだけ確認しましよ (ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速)】 (Plus プランの場合、自動で GPT-4o が選択されます)。生成された画像には透かしなどは表示されませんが、後述するようには見えないメタデータで AI 画像であることがタグ付けされていま (「ChatGPT」、4o モデルによる画像生成が可能に - 無料プランでも - CNET Japan)】。

利用プランと制限: GPT-4o の画像生成機能は無料ユーザーでも利用可能ですが、使用回数に制限がありま (ChatGPT の画像生成 AI 完全ガイド: 無料でも使える最新機能を徹底解説)】。無料プランでは 1 日あたり数回程度の画像生成が可能なのに対し、Plus や Pro など有料プランではほぼ無制限に生成できま (ChatGPT の画像生成 AI 完全ガイド: 無料でも使える最新機能を徹底解説)】。これは従来、画像生成が主に有料ユーザー向け (ChatGPT Plus で DALL-E を利用) だったのに比べ大きな改善で (「ChatGPT」、4o モデルによる画像生成が可能に - 無料プランでも - CNET Japan)】。現在は全ユーザーに段階的に展開中であり、無料プランでも順次使えるようになっていま (「ChatGPT」、4o モデルによる画像生成が可能に - 無料プランでも - CNET Japan) (ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速)】。もし自分のアカウントでまだ GPT-4o が選択できない場合、OpenAI の提供する別アプリ「Sora」経由で先行利用する方法もありま (ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速)】 (Sora では他ユーザーのプロンプト閲覧や Remix 機能による画像部分編集など独自機能も提供されていま (ChatGPT「GPT-4o」に画像生成機能が追

[加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#)) ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))】)。

プロンプト作成のコツ: 基本的に自然文で構いませんが、具体的に描写してほしい要素(場所、対象物、色合い、スタイルなど)を盛り込むとより意図に合った画像が得られます ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))】)。たとえば「公園で遊ぶ子供」よりも「春の公園で桜の木の下ではしゃぐ 2 人の子供(写真風)」のように詳細に書く方が望ましい結果に近づきます。また、一度で完璧な画像を得ようとせず、出力→気になる点を追加指示→再出力という対話サイクルを活用するのがおすすめです ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))】)。こうすることで、モデルの解釈に任せつつ細部だけ調整する効率的な生成が可能です。

コンテンツの安全性と禁止事項: OpenAI はコンテンツポリシーに従い、違反する画像リクエストはブロックしています ([「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#))】)。暴力・ポルノ・ヘイトなど不適切な内容や、政治的ミスインフォメーション、違法行為を助長するような画像の生成指示は受け付けられません。特にヌードや生々しい暴力描写については「以前よりも制限を強化した」としており、安全な応答に細心の注意が払われています ([「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#))】)。一方、従来不可だった有名人の画像生成については、現実の写真を参考に著名人を再現する画像を作れるケースもあるようで ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))】)(※無条件に許可されたわけではなく、公序良俗に反しない範囲で変革・再現する用途に限られると推測されます)。生成された画像には、**C2PA メタデータ**(Content Authenticity Initiative の標準)が自動付与され、ファイル情報から AI 生成と判別できるようになっていま ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))】)。ただし目に見えるウォーターマークは追加されないため、ダウンロードした画像自体には余計なロゴ等は入りませ ([「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#))】)。利用者は、生成された画像であっても事実を示すものではない点に留意し、誤解を招かないよう活用する必要があります。

3. 技術的な仕組み

モデルと技術: GPT-4o は OpenAI が開発したマルチモーダル (omnimodal) 大規模モデルで ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。従来の GPT-4 に画像生成機能を組み込んだような位置付けで、言語と画像の両方をネイティブに扱えるのが特徴で ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。テキストから画像を出力するために、モデル内部に画像生成用のデコーダ (生成ネットワーク) が統合されているとみられます。OpenAI は詳細な技術アーキテクチャを公表していませんが、学術動向から考えて拡散モデル (Diffusion) や Transformer ベースの画像生成ネットワークを GPT 本体に組み込んでいる可能性があります。実際、GPT-4o は画像とテキストの同時分布に基づいて訓練されており、テキストと画像の関係だけでなく画像と画像同士の関連も学習しています ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。その上で ChatGPT のような手法 (人間のフィードバックによる調整など) による事後学習 (Post-training) を施し、視覚的に分かりやすく有用で一貫性のある画像を生成できるよう最適化されています ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。要するに、言語モデルとしての知識・文脈理解力と、画像生成モデルとしての描画力を一体化した「基盤モデル (Foundation Model)」と言えます。

GPT-4o の画像生成は ChatGPT のアーキテクチャ深くに組み込まれているため、テキスト生成と画像生成が密接に連携します ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。これにより、純粋な画像生成モデルでは難しい知識の活用や文脈の保持が可能になります ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。例えば先の例のように、LLM としてユーモアのセンスを活かして画像中のテキスト文言を自動生成した ([OpenAI が画像生成を完璧にした \(比較を含む\) | AGI に仕事を奪われたい](#))、会話の流れに沿って逐次画像を更新したりできます ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。この統合型モデルにより、ChatGPT はテキストと画像を同時に理解・生成できる真のマルチモーダル AI となっています。

性能面の工夫: GPT-4o は、OpenAI が以前提供していた DALL-E 3 シリーズと比べて新規性の高いアプローチを採用しており、著しく強化されています ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。特に、以下の点で技術的進歩がみられます。

- **テキストレンダリングの改善:** モデルが内部で文字列を画像へマッピングする能力を向上させる訓練が行われていま ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。これによりアルファベット以外の言語もある程度正しく描画可能となりました (完全な多言語対応

には今後の改善が予定されています ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。OpenAI は長文の英語テキストを手書きでホワイトボードに書き出すような複雑な例も自然に生成できていると報告しています ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#)) ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))】。

- **詳細プロンプトへの忠実性:** 前述のように多数要素のシーンでも破綻しないよう、モデルの注意機構が強化されています ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。他のシステムでは苦手だった複雑な関係性のレンダリングも、知識と論理推論力を持つ GPT-4o ならではのアプローチで対応しています。
- **知識と推論の直接利用:** 画像生成専用モデルは視覚パターンの学習に特化しますが、GPT-4o ではテキスト知識と視覚を結び付けて学習しているため、**物理法則や世界知識に即した画像**を出力しやすいと言われます ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。例えば「ペンローズの三角形」のような錯視オブジェクトや、「火星に人類が降り立つ未来」のような架空シーンも、言語知識を活かしつつ説得力あるビジュアルに落とし込めます。
- **スタイル多様性:** 学習データに写真、イラスト、絵画、CG など多種多様な画像が含まれており、特定の画風に偏らない出力が得られます ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。Midjourney のようなサービスではデフォルトで独特な芸術スタイルが混ざることがあります ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))、GPT-4o はユーザー指定したスタイルに忠実で、必要に応じてクリエイティブにもリアルにも振れる柔軟さを持ちます ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))】。

安全対策: OpenAI は GPT-4o の開発にあたり強力な安全基準を設けています ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。不適切な画像が生成されにくいよう学習段階でフィルタリングやチューニングが行われており、DALL-E や他モデルの運用から得た教訓が活かされています ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。生成された画像には前述の C2PA メタデータで**出所情報**が埋め込まれ、AI 産品であることの透明性も確保されています ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの](#)

[指定・一貫した画像の複数回出力が可能に | gihyo.jp](#)】。また、「性的なディープフェイク」や「児童虐待コンテンツ」のような悪用リスクの高い生成指示は、検知次第ブロックする仕組みで ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))】。もっとも、ユーザー側でも誤った情報を画像化しないよう注意が必要であり、GPT-4o システムカード追補では新たに生じうるリスクと対策が詳述されています ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))】。

以上のように、「4o Image Generation」は OpenAI 独自の高度なマルチモーダル技術に支えられています。言語モデルと画像モデルのハイブリッドによって、他にはない精度と利便性を実現した点が技術的な大きな強みです。

4. 他の画像生成サービスとの比較

ChatGPT の「4o Image Generation」を、代表的な他の画像生成 AI と機能・品質・自由度・速度・利用条件などで比較します。ここでは、OpenAI の従来モデル DALL-E 3 (ChatGPT 旧画像生成機能の裏側)、Midjourney (高品質画像で知られる商用サービス)、Stable Diffusion (オープンソースの画像生成モデル群) を取り上げます。

各サービスの特徴を以下の表にまとめ、その後に補足解説します。

ChatGPT (4o)	DALL-E 3	Stable Diffusion
フリーアクセス	有料	オープンソース
多様な生成可能	高品質生成	多様な生成可能
生成速度が速い	生成速度が速い	生成速度が遅い
生成画像のクオリティが高い	生成画像のクオリティが高い	生成画像のクオリティが低い
生成画像のスタイルが豊富	生成画像のスタイルが豊富	生成画像のスタイルが豊富
生成画像の用途が広い	生成画像の用途が広い	生成画像の用途が広い
生成画像の著作権が不明	生成画像の著作権が不明	生成画像の著作権が不明
生成画像のセキュリティが低い	生成画像のセキュリティが低い	生成画像のセキュリティが低い
生成画像のプライバシーが低い	生成画像のプライバシーが低い	生成画像のプライバシーが低い
生成画像のセキュリティが高い	生成画像のセキュリティが高い	生成画像のセキュリティが高い
生成画像のプライバシーが高い	生成画像のプライバシーが高い	生成画像のプライバシーが高い

Ch
at
GP
T(
E GP
T-
4o)

ラ
ス
ト
ま
で
高
品
([A
dde
ndu
m
to
GP
T-
4o
Sy
ste
m
Car
d:
4o
ima
ge
gen
era
tio
n

DALL-E 3

St
ab
le
Di
ff
us
io
n
現
た
に
ビ
優
ジ
れ
ユ
る
ア
ル
写
。
真
デ
的
フ
な
オ
ル
描
写
ト
も
で
可
ド
能
ラ
だ
マ
が
チ
ー
ツ
部
ク
不
な
自
作
然
風
さ
。
も
。

ChatGPT (GPT-4o)

非常に高い。複雑な指示も厳密に再現(Ch atGPT「GPT-4o」に画像生成

DALL-E 3

高いが GPT-4o には劣る。複数要素の細部で崩れる場合も。

Stable Diffusion モデルとプロンプト技術次第。思い通りにするには試行錯誤
Midjourney 中程度。細部指定は抜け落ちがち。概略を雰囲気良く表現する

Chat
at
GP
T((E
GP
T-
4o)

[機能](#)
[が](#)
[追](#)
[加](#)
[！](#)
[特](#)
[徴](#)
[や](#)
[使](#)
[い](#)
[方](#)
[を](#)
[徹](#)
[底](#)
[解](#)
[説](#)
[↓](#)
[SHI](#)
[FT](#)
[AI](#)
[TI](#)
[ME](#)
[S](#)
[↓](#)
[AI](#)
[の](#)
[情](#)

DALL-E 3

Stable
Diffusion
Milestone
journey
傾向。
。

Ch
at
GP
T(
E GP
T-
4o)

報
イ
ン
フ
ラ
を
構
築
し
、
日
本
の
AI
推
進
を
加
速
)
(Ch
atG
PT
、
実
在
人
物
の
写

DALL-E 3

St
ab
le
Mi
dj
ou
rn
ey
Di
ff
us
io
n

Chat
GPT
T
GPT-
4o)

真生成など画像生成機能強化

PC Watch】。

段落レベルの長文も

DALL-E 3

短い単語は概ね可。長文や日本語は崩れやすい。

Stable
Diffusion

基本苦手。アナル。専用モデル

Chat
at
GP
T
E
GP
T-
4o)

正
確
に
描
([OpenAI](#)
[が](#)
[画](#)
[像](#)
[生](#)
[成](#)
[を](#)
[完](#)
[璧](#)
[に](#)
[し](#)
[た](#)
([比](#)
[較](#)
[を](#)
[含](#)
[む](#))
|
[AGI](#)
[に](#)
[仕](#)
[事](#)

DALL-E 3

Stable
Diffusion
journey
ion
n
トル
数や
後後
文後
字処
程理
度で
な補
らう
稀必
に要
可。可
。文
章は
まず
不
可
能。
。

ChatGPT (GPT-4o)

を奪われたたい

】。日本語含め短文ならばほぼ完璧。

チャットで何度

DALL-E 3

ChatGPT 統合では一枚生成ごとに再プロンプト必 (「[ChatGPT](#)」、[4o モデルによる画像生成が可能に--無料プランでも - CNET Japan](#))】。直接の部分編集機能は無し。

Stable Diffusion
Midjourney

専用UIでは不便

Ch
at
GP
T(
GP
T-
4o)

も
編
集
可
([Ch
atG
PT
「G
PT
=
4o」](#)
に
[画
像
生
成
機
能
が
追
加
！
特
徴
や
使
い
方](#)

DALL-E 3

St
ab
le
Di
ff
us
io
n
可
ペ
。
イ
ン
V
ari
ト
等
io
ns
機
能
が
対
話
的
生
成
は
い
可
。
手
動
操
作
が
指
定
で
き
ず
。

Chat
at
GP
T(GP
T-
4o)

[を徹底解説 | SHI FT AI TI ME S | AI の情報インフラを構築し、日本の](#)

DALL-E 3

Stable
Diffusion
Midjourney

Chat
GPT
T
GPT-
4o)

[AI 推進を加速](#)

】。画像の一部変更も指示で実行。

画像アップロード

(ChatGPT では非対応※)元画像のバリエーション生成は可能。

DALL-E 3

Stable
Diffusion

画像
プロンプト
画像
機

Chat
at
GP
T(
E
GP
T-
4o)

一
ド
に
対
応
。
内
容
解
析
し
て
変
換
・
続
き
生
成
が
可
能
。
([ChatGPT](#)、[実在人物](#)

DALL-E 3

St
ab
le
Di
ff
us
io
n
ト能
指、
定C
可on
(tr
sol
タN
イet
ル等
反豊
映富
程。高
度)高
度な
直変
接換
のも
変設
換定
は次
第困
難。
。

Chat
at
GP
T(
E GP
T-
4o)

DALL-E 3

St
able
Mid
jour
ney
Diff
usion
n

の
写
真
生
成
な
ど
画
像
生
成
機
能
強
化

ニ
PC
Wa
tch
】。

指
定
し
た
通
り
の

幅広いが、ややイラスト的な出力傾向。

多
彩
デ
ル
が
毎
常
に
専
門

Ch
at
GP
T(
E GP
T-
4o)

作
↑風
に
柔
軟
に
対
([Ch](#)
[atG](#)
[PT](#)
[「G](#)
[PT](#)
[=](#)
[4o」](#)
に
[画](#)
[像](#)
[生](#)
[成](#)
[機](#)
[能](#)
が
[追](#)
[加](#)
[!](#)
[特](#)
[徴](#)
[や](#)

DALL-E 3

St
ab
le
Mi
dj
ou
rn
ey
Di
ff
us
io
n
ス
タ
イ
ル
な
い
ル
有
気
。
モ
デ
ル
を
切
り
替
え
れ
ば
自
由
。
St
able
Diffusion
の
ス
タ
イ
ル
な
い
ル
有
気
。
モ
デ
ル
を
切
り
替
え
れ
ば
自
由
。
Stable Diffusion
の
画
像
生
成
機
能
が
追
加
！
特
徴
や

Chat
GP
T(GP
T-
4o)

[使い方を徹底解説 | SHI FT AI TI ME S | AI の情報インフラを構築し、](#)

DALL-E 3

St
ab
le
Di
ff
us
io
n
Mi
dj
ou
rn
ey

[較を含む | A GI に仕事を奪われたい |](#)
。

ChatGPT (GPT-4o)

[日本のAI推進を加速](#)】。写真・絵画など幅広い。

ChatGPT
・上
で
対
話

DALL-E 3

Stable Diffusion

専用
GUI
で
スク

ChatGPT で指示するだけ(Plus 限定でした)。

Ch
at
GP
T(PT
GP
T-
4o)

するだけ。
日本語OKで初心者向け
([ChatGPTの画像生成AI完全ガイド](#)):

DALL-E 3

Stable
Diffusion
マリンプロトタイプ使用
。高度な知識が要求される場合も。
英語推奨でやってみよう。

Chat
GP
T(GP
T-
4o)

DALL-E 3

St
ab
le
Di
ff
us
io
n

無料で使える最新機能を徹底解説)】。

1 枚 [gihyo.jp\]\(https://gihyo.jp/article/2025/03/gpt4o-image-generation#:~:text=%E4%BB%8A%E5%9B%9E%E3%81%AE%E3%83%A2%E3%83%AB%E3%82%92%E4%BD%BF%E3%81%A3%E3%81%9F%E7%94%BB%E5%83%8F%E7%94%9F%E6%88%90%E3%81%AF%E5%A4%9A%E3%81%8F%E3%81%AE%E5%A0%B4%E5%90%88%E3%80%81%E6%9C%80%E5%A4%A7%E5%88%86%E3%81%8B%E3%81%8B%E3%82%8B%E3%81%A8%E3%81%84%E3%81%86%E3%80%82%E3%81%AA%E3%81%8A%E3%80%81%E3%82%A2%E3%82%B9%E3%83%9A%E3%82%AF%E3%83%88%E6%AF%94%E3%80%81%16%E9%80%B2%E3%82%B3%E3%83%BC%E3%83%89%E3%82%92%E4%BD%BF%E7%94%A8%E3%81%97%E3%81%9F%E6%AD%A3%E7%A2%BA%E3%81%AA%E8%89%B2%E3%80%81%E](https://gihyo.jp/article/2025/03/gpt4o-image-generation#:~:text=%E4%BB%8A%E5%9B%9E%E3%81%AE%E3%83%A2%E3%83%AB%E3%82%92%E4%BD%BF%E3%81%A3%E3%81%9F%E7%94%BB%E5%83%8F%E7%94%9F%E6%88%90%E3%81%AF%E5%A4%9A%E3%81%8F%E3%81%AE%E5%A0%B4%E5%90%88%E3%80%81%E6%9C%80%E5%A4%A7%E5%88%86%E3%81%8B%E3%81%8B%E3%82%8B%E3%81%A8%E3%81%84%E3%81%86%E3%80%82%E3%81%AA%E3%81%8A%E3%80%81%E3%82%A2%E3%82%B9%E3%83%9A%E3%82%AF%E3%83%88%E6%AF%94%E3%80%81%16%E9%80%B2%E3%82%B3%E3%83%BC%E3%83%89%E3%82%92%E4%BD%BF%E7%94%A8%E3%81%97%E3%81%9F%E6%AD%A3%E7%A2%BA%E3%81%AA%E8%89%B2%E3%80%81%E) 1 枚 10 枚 ~ 20 15 月 ~ 秒 30 程 秒 度 程で 度 高 速 複。

ChatGPT (GP T-4o) に組み込まれた画像生成機能を提供開始—画像内テキ

DALL-E 3

、 [9%80%8F%E6%98%8E%E3%81%AA%E8%83%8C%E6%99%AF%E3%81%AA%E3%81%A9%E3%82%82%E6%8C%87%E7%A4%BA%E3%81%A7%E3%81%8D%E3%82%8B%E3%80%82\)\)】](#)。順次表示される。

Stable Diffusion 標準出力4枚も一度に生成でき・4枚選択)。

Chat
at
GP
T((GP
E T-
4o)

ストの指定一貫した画像の複数回出力が可能に

厳格なポリシー

DALL-E 3

厳しい制限あり。有名人や過激表現は不可。

Stable
Diffusion
Midjourney

制限なし
制限あり
自

Chat
GP
T(GP
T-
4o)

一
で
不適切
内容は
不
([Chat](#)
[GP](#)
[T](#)
[4o](#)
モデル
による
画像
生成
が
可能

DALL-E 3

Stable
Diffusion
Midjourney
り己
(責任
露骨)
な。
性的
公開
・サ
暴一
力ビ
ス上
では
禁止
各社
規制
あり
のり
写。
真風
NG
。

Chat
at
GP
T((E
GP
T-
4o)

DALL-E 3

St
ab
le
Mi
dj
ou
rn
ey
Di
ff
us
io
n

[に-](#)
[-無](#)
[料](#)
[プ](#)
[ラ](#)
[ン](#)
[で](#)
[も](#)
[ニ](#)
[CN](#)
[ET](#)
[Ja](#)
[pan](#)
)]。有名人画像は条件付き許可。

Ch
at
GP
T(
E GP
T-
4o)

可
能
(利
用
規
約
内
で
([Ch](#)
[atG](#)
[PT](#)
[「G](#)
[PT](#)

可
能
に

[画](#)
[像](#)
[生](#)
[成](#)
[機](#)
[能](#)
[が](#)
[追](#)
[加](#)
[!](#)
[特](#)
[徴](#)

可能。生成物の権利はユーザーに帰属。

DALL-E 3

St
ab
le
Mi
dj
ou
rn
ey
Di
ff
us
io
n

有
料
可
能
会
員
(
は
モ
商
デ
ル
利
に
用
よ
り
。異
(
な
る
成
が
物
基
の
本
権
制
限
を
な
し
張
)
可
。

Chat
at
GP
T(GP
T-
4o)

[や
使
い
方
を
徹
底
解
説
|
SHI
FT
AI
TI
ME
S
|
AI
の
情
報
イ
ン
フ
ラ
を
構
築](#)

DALL-E 3

St
ab
le
Di
ff
us
io
n
Mi
dj
ou
rn
ey

ChatGPT (GPT-4o)

[し、日本のAI推進を加速](#)】。著作権はユーザーに帰属。

DALL-E 3

Stable Diffusion

※ChatGPT 経由の DALL-E 3 では画像アップロードを用いた編集はできませんでした (DALL-E 2 では専用エディタで対応)。

上記の比較から、GPT-4o 搭載の ChatGPT は総合的に優れ、実用性が非常に高いことがわかります。他サービスと比べた際の注目ポイントを補足します。

- **プロンプト遵守とテキスト描画の正確さ**において、GPT-4o は突出しています。他モデルが苦手とする長い文章の埋め込みも正確で、チケットのように複数段落の文字情報を含む画像さえ破綻なく生成できま ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。Midjourney や Stable Diffusion ではまず困難な芸当であり、Ideogram といった一部の特殊モデルを除けば GPT-4o が頭一つ抜けていま ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。また、GPT-4o は指示された構図・内容をそのまま反映する忠実性も極めて高 ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))、ユーザーの細かな要望(例:「A と B が握手している横で C が拍手している」等)を理解して表現できます。他サービスではこのレベルの細密なプロンプト解釈は難しく、Midjourney は芸術性は高いものの細部指定をしばしば無視しがちという指摘がありま ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。
- **対話的な画像編集機能**は ChatGPT (GPT-4o)の大きな強みです。他の画像生成サービスでは、一度生成した画像に後から変更を加えたい場合、再度プロンプトを調整して一から生成し直す必要がありました。Midjourney にはバリエーション生成やアップスケール機能はありますが、「先ほどの画像の背景だけ夜に変える」といった部分的な修正はできませ ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。Stable Diffusion は高度なユーザーであれば別途インペイントツール等で可能ですが、対話形式で手軽に…とはいきませ ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。GPT-4o は会話の文脈を保持したまま画像を更新・編集できるため、まさに Photoshop を AI アシスタントに任せるような使い勝手で ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#))。この編集容易性は、クリエイティブ作業のワークフローを大きく効率化すると期待されています。
- **コンテンツの自由度**に関しては、Stable Diffusion(ローカル版)のような無制限には及ばないものの、GPT-4o も従来モデルに比べ柔軟になっています。例えば、有名人の写真風画像は DALL-E 3 では厳禁でしたが、GPT-4o ではユーザーが提供した既存の写真を元にした変換であれば生成可能とされていま ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。完全な架空捏造は不可でも、肖像画風イラスト程度なら許容されるケースが増

えたようです。ただしポリシー違反の境界は慎重に設定されており、過激なポルノ・暴力は依然として厳しく禁止されています（[「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#)）。一方で Stable Diffusion は完全オープンな分リスクも伴うため、商用プロジェクトなどでは信頼性の面で GPT-4o や Midjourney が選好される場面も多いでしょう。

- **出力品質とスタイル**では、Midjourney が依然トップクラスとの評価もあります。Midjourney はプロンプトが多少粗くとも非常に洗練されたビジュアルを生成し、「クリエイティブディレクターが仕上げたような芸術性」が感じられると評されま（[OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われたい](#)）。GPT-4o も美しい画像を作れますが、Midjourney 独特の絵画的表現を望む場合は好みが分かれるかもしれません。ただし GPT-4o はユーザーの指示次第で Midjourney 風の作風も再現可能であり、デフォルトのクセが強くない分、用途に合わせた柔軟な表現ができます（[ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#)）。例えばビジネス用途でインフォグラフィックを作る際、余計な芸術的アレンジをせず指示通りの図表を描いてくれるのは GPT-4o の利点で（[「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#)）。
- **利用の手軽さとコスト**では、ChatGPT (GPT-4o)は無料プランでも使える点で優れています（[「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#)）。Midjourney は現在無料利用が原則できず、有料サブスクリプションが必要です。一方、ChatGPT は回数制限はあるものの**無料枠**がありますし、有料プランでも月額 20ドル程度で他の高度な言語生成機能も含め利用できます。Stable Diffusion はソフト自体は無料ですが、環境構築や GPU 資源のコストが発生します。その意味で、**総合的なコストパフォーマンス**も ChatGPT の画像生成機能は高いと言えるでしょう。

以上を踏まえると、ChatGPT の「4o Image Generation」は**汎用性・正確性・利便性のバランスが取れたサービス**です。他のサービスと比べ、「ここが突出している」「ここは劣る」というより、幅広いニーズに平均点以上で応えられる点が強みです。特にテキストと画像を組み合わせた応用（資料作成、プレゼン図作りなど）では唯一無二の存在であり、従来別々だった作業を一元化できる革新的なツールとなっています。

5. リリース日とアップデート履歴

リリース日: ChatGPT の画像生成機能「4o Image Generation」は、**2025 年 3 月 25 日 (米国時間)**に公開が発表され、日本時間では 3 月 26 日に提供開始されました ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。OpenAI の公式 Twitter でも「4o Image Generation が到来した。本日より Plus、Pro、Team、そして Free ユーザーに順次ロールアウトする」と告知されています ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。このアップデートにより、無料含む全ユーザーが ChatGPT 内で画像をネイティブに生成できるようになりました ([「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#))。

主なアップデートの経緯:

- **2023 年 9 月～10 月:** OpenAI は画像生成モデル DALL-E 3 を発表し、ChatGPT Plus/Enterprise ユーザー向けにプレビュー提供を開始しました (同年末には無料ユーザーも 1 日 2 枚まで DALL-E 3 を利用可能に ([ChatGPT DALL-E 3: Complete Guide \(Generate Images with Text\)](#)) ([OpenAI allowing ChatGPT users to create two DALL-E 3 images for ...](#)))。これが ChatGPT への画像生成機能の初導入で、当初は GPT-4 が裏で DALL-E API を呼び出す形で実現されていました。
- **2024 年:** ChatGPT + DALL-E 3 による画像生成が広く普及。ユーザーからは高度なアート表現力が評価される一方、文字生成の不得手さや詳細指示の限界など課題も指摘されました。OpenAI はこれら課題解決に向け、言語モデルと画像モデルの統合研究を進めていきま ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。一方で、他社からも Midjourney v5 の改良版や新サービス (例: Ideogram のテキスト特化 AI) などが登場し、画像生成 AI 競争が激化した時期です。
- **2025 年 3 月: GPT-4o モデル登場。** OpenAI は GPT-4 の後継となるマルチモーダルモデル「GPT-4o」にネイティブ画像生成機能を統合し、ChatGPT および実験的動画 AI の Sora で提供開始しまし ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#)) ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。このアップデートは単なる機能追加に留まらず、ChatGPT の基盤モデル自体を入れ替える大規模なものでし ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。テキストレンダリング精度、複雑指示への対応力、マルチターン編集など多くの面で飛躍的進化を遂げていま ([ChatGPT『GPT-4o』の画像生成](#)

[が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#)) ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。リリース直後から「過去最強の画像生成 AI」と国内外メディアで話題となりました。

- **2025年4月(予定):** 開発者向けの画像生成 API が提供予定とアナウンスされています ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。これにより ChatGPT 以外のアプリケーションやサービスでも GPT-4o の画像生成機能を統合できるようになります。公開時点では ChatGPT と Sora での利用のみでしたが、数週間以内に API 経由でも利用可能になる見込みで ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。また、同時期にモデルの細かな不具合修正も進められ、アップロード画像の顔編集の一貫性バグ修正 (リリース 1 週間以内に対応) や、画像編集時の精度向上などが順次行われていま ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。
- **今後の展望:** GPT-4o は今後も継続アップデートが予定されています。現時点での課題として、非ラテン文字 (日本語以外の多言語) のテキスト描画精度や、グラフ生成の正確さ、極端に多数要素を含む画像 (例えば完全な周期表の図など) のレンダリングといった点が挙げられていま ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))。OpenAI はこれらの精度向上に取り組んでおり、今後のアップデートでさらに改善される可能性があります ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。特に日本語を含む多言語対応は、デザインや教材作成への活用範囲を広げるため期待されていま ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。

さらに、GPT-4o は「Omnimodal (全てのモーダルに対応する)」モデルと位置づけられており、将来的には動画生成や音声出力などへの発展も示唆されています ([Addendum to GPT-4o System Card: 4o image generation | OpenAI](#))。実際、OpenAI は GPT-4o の技術を用いた実験的な動画生成 AI「Sora」を公開してお ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫](#)

[した画像の複数回出力が可能に | gihyo.jp](#)】、画像のみならず動画や 3D といった次の領域も視野に入れているようです。**マルチモーダル AI の基盤**として GPT-4o を磨き上げ、将来的な GPT-5 などより高度なモデル開発につなげる狙いもあるでしょう。

総じて、2025 年 3 月の GPT-4o 画像生成機能のリリースは ChatGPT の進化におけるマイルストーンとなりました。今後はユーザーからのフィードバックを踏まえて細部が改良され、API 提供によるエコシステム拡大や、更なるモーダル統合が進むと予想されます。

6. 従来の画像生成機能との違い

最後に、ChatGPT に以前から搭載されていた画像生成 (DALL-E ベース) と「4o Image Generation」との**進化ポイント**を整理します。

- **モデル統合の深さ:** 最大の違いは、画像生成が言語モデルにネイティブ統合されたことで ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。従来は ChatGPT から DALL-E という外部の画像モデルを呼び出す形で、「テキストは GPT-4、画像は DALL-E」と役割分担していました。アップデート後は GPT-4o モデルがテキストも画像も一手に生成できるようになり、シームレスな動作を実現しています ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。これにより、画像生成も会話履歴や文脈を完全に共有した上で行われ、テキスト応答と画像応答の間に知識・情報のギャップが無くなりまし ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。例えば会話で取得したデータを使って即座に図解する、といったこともモデル内で完結します。
- **生成内容の正確さ向上:** テキスト指示に対する忠実度が格段に上がりました。テキスト内の文字描画 ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))、複雑なシーンの再現 ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#)) など、DALL-E 3 では不十分だった点が大きく改善されています。特に日本語テキストの描画は、旧 DALL-E ではほぼ不可能だったのが、GPT-4o では短い文であれば正しく再現できるまでになりました ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。また、一度に扱える要素数も増え、**実用的な図表や長文入り画像**などが生成可能になったのは大きな進歩で ([ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch](#))。

- **マルチターンの編集機能:** 従来の ChatGPT+DALL-E 連携では、画像を出力した後にさらに編集したい場合、ユーザーがプロンプトを書き直して新たに画像生成をやり直す必要がありました。過去の画像との一貫性を保つ仕組みはなく、**逐次的なデザイン調整**がしにくかったのです。GPT-4o ではこの弱点が解消され、**チャット形式での画像編集**が正式に導入されました ([「ChatGPT」、4o モデルによる画像生成が可能に--無料プランでも - CNET Japan](#))。例えば「もう少し笑顔にして」「背景に看板を追加して」と対話を重ねながら、**同じ登場人物・スタイルを維持して画像をアップデート**できます ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))。これは従来 DALL-E シリーズには無かった柔軟性であり、クリエイティブワークの効率と自由度を飛躍的に高めています。
- **UI 上の違い:** 利用インターフェース自体は大きく変わりませんが、**モデル選択**に GPT-4o が加わった点の変化です。以前は ChatGPT で画像生成する際、「Vision」機能を ON にした GPT-4 が裏で DALL-E を使う形でした。アップデート後はモデル一覧に「GPT-4o」が登場し、このモードでテキストも画像も生成するようになっていま ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。また SHIFT 社のレポートによれば、ChatGPT 内で**「DALL-E GPT」**という旧モデルも引き続き選択可能で、必要に応じて従来の DALL-E 3 による画像生成も使えるようで ([ChatGPT「GPT-4o」に画像生成機能が追加！特徴や使い方を徹底解説 | SHIFT AI TIMES | AI の情報インフラを構築し、日本の AI 推進を加速](#))。ただし基本的には GPT-4o の方が高性能なため、特別な理由が無ければ新機能を使うメリットが大きいです。
- **速度・コストの変化:** 画像生成に要する時間は GPT-4o の方が若干長めです (数十秒~1 分程度)。DALL-E 3 も高速ではありませんでしたが、GPT-4o は大規模モデル統合の影響で DALL-E よりやや**処理に時間がかかる**との指摘があります ([OpenAI が画像生成を完璧にした\(比較を含む\) | AGI に仕事を奪われない](#))。しかし、その分**無料ユーザーにも開放された価値は大きく**、Plus 限定だった DALL-E 統合に比べて裾野が広がりました ([「ChatGPT」、4o モデルによる画像生成が可能に--無料プランでも - CNET Japan](#))。有料プランなら従来通り追加費用なく使い放題であり、画像生成ごとに別料金が発生するようなことはありません (※ChatGPT の定額課金に含まれる範囲です)。
- **品質と安全性:** 出力される画像の品質は総じて向上しました。特に**文字入り画像や実用図解**といった用途では、DALL-E 搭載時代から飛躍的に実用レベルに達していま ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#)) ([ChatGPT『GPT-4o』の画像生成が大幅ア](#)

[アップデート！新機能を徹底解説 | ChatGPT 研究所](#)】。一方、[安全性ポリシーの強化も行われていま \(「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan\)](#)】。これは新機能というより運用面のアップデートですが、前述のようにヌーディティや流血表現等は以前より厳しくブロックされる傾向がありま ([「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan](#))】。また、生成画像へのメタデータ付与 (C2PA) は DALL-E 3 の後期から導入されていましたが、GPT-4o でも継続して適用されています ([OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp](#))】。

まとめると、「4o Image Generation」は ChatGPT の画像生成能力を質・量ともに大きく底上げしました。従来の DALL-E 連携では成し得なかったテキストと画像の真の融合* ([ChatGPT『GPT-4o』の画像生成が大幅アップデート！新機能を徹底解説 | ChatGPT 研究所](#))】が遂げられた点で、単なる進化ではなく次元の異なるアップデートと言えます。ユーザー目線では、「できること」が増えただけでなく「使い勝手」も格段に良くなっています。今後この差はさらに開いていくと見られ、ChatGPT は統合 AI プラットフォーム**として他の画像生成サービスとの差別化を一層強めていくでしょう。

参考文献・情報源:

- OpenAI, *Addendum to GPT-4o System Card: 4o image generation*, 202 (Addendum to GPT-4o System Card: 4o image generation | OpenAI) (Addendum to GPT-4o System Card: 4o image generation | OpenAI) (Addendum to GPT-4o System Card: 4o image generation | OpenAI)】
- PC Watch, 「ChatGPT、実在人物の写真生成など画像生成機能強化」(2025年3月26日 (ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch) (ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch) (ChatGPT、実在人物の写真生成など画像生成機能強化 - PC Watch)】
- CNET Japan, 「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも(2025年3月26日 (「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan) (「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan) (「ChatGPT」、4o モデルによる画像生成が可能に—無料プランでも - CNET Japan)】
- gihyo.jp, *OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始*(2025年3月26日 (OpenAI、GPT-4o に組み込まれた画像生成機能を提供開始 —画像内テキストの指定・一貫した画像の複数回出力が可能に | gihyo.jp)

