

生成AIによるR&D業務の効率化と意思決定へのシフト

近年の生成AI（Generative AI）は急速に進化し、2025年12月にはOpenAIのGPT-5.2やGoogle DeepMindのGemini 3 Deep Thinkといった**第3～4世代の最新モデル**が登場しました¹²。従来、生成AIには誤情報ルシネーション）を生成する、不確実な推論や創造性の限界がある、応答の再現性が低く説明責任（説明性）の欠如がある等の批判がありました。しかし、最新モデルは大規模データと高度なアーキテクチャにより**汎用知能や長文理解、マルチモーダル処理**の性能が飛躍的に向上し、企業利用や研究開発への適用が加速しています³⁴。本レポートでは、GPT-5.2やGemini 3 Deep Thinkを例に、新世代AIモデルがこうした否定的見解をどのように克服しつつあるかを整理し、特に**創薬、材料開発、製造業の設計・試作**の分野で生まれた成果や発明、プロセス革新を概観します。併せて、依然残る課題や専門家による慎重な見方についても述べます。

最新モデルによる課題克服のポイント

誤情報・ハルシネーション対策の進展

最新の生成AIモデルは**事実に反する出力（ハルシネーション）の低減**に取り組んでいます。GPT-5.2ではモデルを「**検索ツール**」やコード実行と組み合わせて推論させることで事実性が向上し、ChatGPT上のテストで**誤答含有率が前世代比30%減少**したと報告されています⁵。これは専門家による調査や分析支援において、より**信頼性の高い回答**が得られることを意味します⁵。また、OpenAIは「重要な問いでは依然としてモデルの回答を人間が検証すべきだ」と注意喚起しつつ、さらなる誤情報抑制に向けた研究を続けています⁶⁷。一方、GoogleもGemini 3を企業向けに提供する際、「当社で**最も事実に忠実なモデル**」であると強調し、大量のテキストや画像・動画データを横断分析することで自信を持って業務知見を引き出せるとしています⁸。Gemini 3の高精度版「Deep Think」モードでは、**回答までにより長い思考時間を割り当てて自己検証**する仕組みを導入し、他モデルでは解けない複雑な論理問題にも正答率を向上させました⁹。例えば人類の知識テスト「Humanity's Last Exam」でツール未使用でも41%という業界最高の正答率を達成しており、**テスト時に計算資源を多く使い推論を吟味**することで回答精度を高めています¹⁰⁹。総じて、新世代モデルは**内蔵の事実知識の精度向上**や**外部ツール連携・推論過程の強化**によって、誤情報生成という否定的要素の克服に大きく前進しています。

創造性・問題解決力の飛躍

創造性の限界は従来モデルの弱点とされましたが、GPT-5.2やGemini 3では**未知の問題への対応力や独創的なアイデア創出**が向上しています。Gemini 3 Deep Thinkモードは「**創造性や戦略的思考を要する課題**」を従来以上に解決できるよう設計されており、段階的な改善を伴う複雑な問題に取り組む能力を備えています¹¹。実際、最新モデルはこれまで人間研究者でも難航していた課題に新提案を行う事例が増えています。例えばGPT-5は、数学の未解決問題（Paul Erdős提起の数論問題）の証明において「ある奇数がパターンを崩す」という**新しいアイデアを提示**し、研究者の行き詰まりを打開しました¹²。このように**異分野の知識を結び付けた直感的発想**は従来モデルには難しく、最新AIの創造性が飛躍している証拠といえます。さらに、GPT-5シリーズは大幅に長い文脈を保持できるため（数十万トークン規模）、大量の論文やデータを俯瞰しながら**新たな仮説やデザイン案を生み出す**ことが可能になりました¹³¹⁴。実際、OpenAIの研究チームによれば、GPT-5は数学・物理・生物など複数領域の知見を組み合わせ**て人間の発想を超える洞察**を示すケースが出現しており、科学者との対話を通じて**概念上の文献探索や仮説提案**など創造的プロセスを支援しています¹⁵¹⁶。Gemini 3もまたマルチモーダル（テキスト・画像・音声・コード）対応力を活かし、例えばUIデザイン分野では1つのプロンプトから**多彩なワイヤーフレームや高精細プロトタイプを自動生成**する柔軟性を示しており、「熟練のデザイナー」のような創造力を発揮できるとの評価があります¹⁷¹⁸。このように、**大規**

模知識+高度推論の最新AIは単なる模倣を超えて新機軸のアイデアを提示し、創造性の限界という批判を着実に克服しつつあります。

再現性・説明性（説明責任）の向上

生成AIの**出力の再現性**や**説明可能性**も改善が進んでいます。以前は同一の質問でも回答が安定しないことが課題でしたが、最新モデルでは**出力の一貫性向上**に向けた工夫が凝らされています。例えばGPT-5.2は内部で「思考プロセス」を扱う**チェーン・オブ・ソート（思考の連鎖）**を最大限に活用し、問題解決手順を忠実に踏むよう最適化されています¹⁹⁷。さらにモデル自身がツールを呼び出し中間結果を検証することで、**常に同じ手順を踏んだ場合は同じ結論に至る**再現性が高まるといいます。説明性についても、OpenAIは研究者と協力しモデルの推論過程を人間が**モニター・検証できる技術**を模索中です⁷。実際の研究例では、GPT-5が難解な数学問題の証明を与えられた際、**途中経過の論理を詳細に示す**ことで人間研究者が検証・指摘を行えるケースも報告されています²⁰²¹。ただし現時点でもモデルはしばしば**もっともらしいが誤った引用やメカニズム**を作り出すことがあり⁷、**出典の明示や根拠の説明**には課題を残します。例えばGPT-5がある数学定理の証明を導いた際、実は過去に発表済みの証明と本質的に同じ内容を再現していたものの、初回の回答では**既存文献への言及がなく**（新規に自力で発見したように見えた）という出来事がありました²²。後から新しいセッションで問うとようやく先行研究を指摘したため、**AIの推論過程への適切な引用付与や説明責任**の難しさが浮き彫りになったとされています²²。この点について専門家は、「**モデルが生成したアイデアの出典や根拠を人間がチェックし補完することが依然重要だ**」と強調しています²³。とはいえ、企業向けにはモデルの判断根拠を可視化する試みも始まっており、例えばGoogleはGemini 3を用いた業務システムで重要な決定プロセスにおいて**AIの推論ログを監査できる機能**の開発を進めています（安全性セクションより）²⁴²⁵。総じて、完全な透明性の実現には尚時間を要しますが、**最新世代モデルは再現性・説明性の面でも改善を重ねており**、ユーザがモデルの思考を追跡・理解しやすくする工夫が進んでいます。

創薬分野への応用と成果

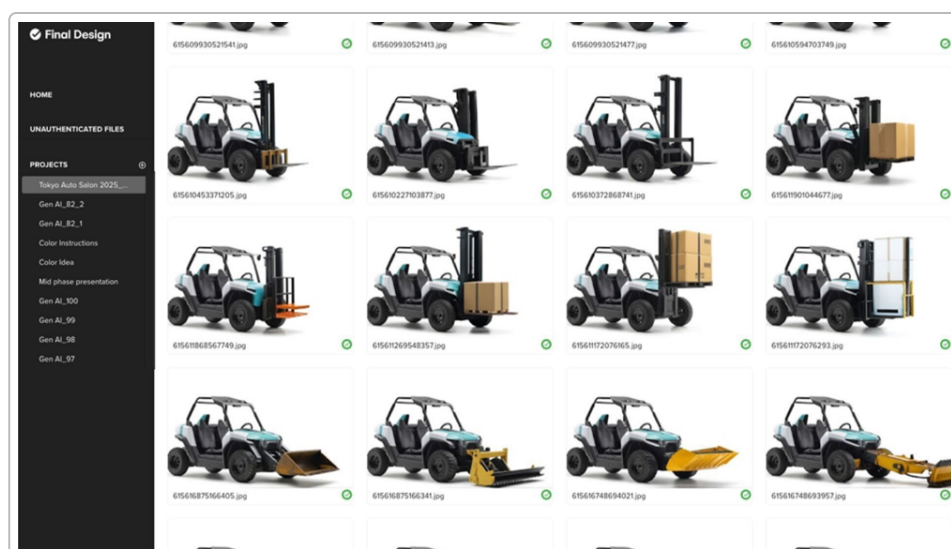
創薬の領域では、GPT-5.2やGemini 3といった強力な生成AIが**新薬候補の発見や開発プロセスの効率化**に寄与し始めています。OpenAIのGPT-5はヘルスケア応用に特化した調整も行われ、「医療博士号を持つ専門家」に匹敵する知識推論能力を備えたと評されています²⁶。実際、バイオ医薬企業のAmgen社はGPT-5をいち早く試験利用し、**複雑な創薬データの深い推論に優れることを確認**しました²⁷。OpenAIの発表によれば、GPT-5は科学文献や臨床データの分析を通じて創薬ターゲットの発見に有用であり、社内テストでも**複雑な医療ポリシーと患者症例のマッピング**など臨床推論で最高性能を示したといいます²⁸。具体的な成果例としては、免疫学者のDr.オーラル・アルパン氏が**GPT-5 Proを用いたドラッグリポジショニング（既存薬の新適応探索）**で顕著な成果を報告しています²⁹³⁰。ある難治性の食物アレルギー患者の症例を入力すると、GPT-5は既知の治療薬リストから**有力な薬剤候補としてデュピルマブ（Dupilumab）を最上位に挙げ**、皮膚疾患への適応から着想を得て「腸の粘膜免疫への作用」という仮説を提示しました³¹³²。この予測は偶然にもアルパン氏らが臨床観察で掴んでいた所見と一致し、さらにモデルは**当該薬の副作用リスクを抱えやすい患者プロファイル**にも言及するなど、人間の科学的直観に近い洞察を示しています³²。アルパン氏は「科学的センスの意味でAIが直感的に振る舞える証拠を見た」と評価し、現在GPT-5の助言を得ながら未治療の難病に対する新たな候補薬を複数ピックアップしています³³³⁴。例えば喘息薬として開発中止となった**Fevipirant**に、自律神経系の難病POTS（体位性頻脈症候群）への有効性を見出すなど、GPT-5が**異領域の知識を結合して提案した新作用機序の仮説**が生まれています³⁴。これらは現時点で仮説段階ながら、AIがヒントを与えたアイデアを元に臨床試験が計画される動きもあり、**新世代AIが創薬のスピードと射程を拡大していることは確実**です。さらに製薬大手やヘルステック企業も続々と最新モデルを導入しており、GPT-5のAPIは米国連邦政府の全職員にも提供されるなど³⁵、**創薬研究から臨床支援まで幅広く生成AIが実装段階に入っています**。一方で、医療領域では誤情報のリスクが直結して患者の生死に関わるため、専門家は「AIが患者情報を**幻覚（でっち上げ）**した場合、それが重要情報の見落としや虚偽の病歴に繋がらう」と警鐘を鳴らしています³⁶。実際、GPT-5.1までは医療QAで依然ハルシネーションが見られたとの指摘もあり、最新モデルを活用する際には**医師や研究者による検証・補完が不可欠**との認識が共有されています³⁶。総じて、創薬

分野では最新AIが**創造的な仮説提案や膨大な知識処理で革新をもたらしつつも、最終的な科学的妥当性の判断には人間の目が必要**というバランスで実用が進んでいます。

材料開発（マテリアルサイエンス）への応用と成果

材料科学の領域でも、生成AIは**新材料の探索・設計**に革命を起こしつつあります。特に大規模言語モデルと科学データベースを組み合わせることで、**これまで発見されていない物質や材料特性の予測**を高速化する試みが進んでいます。ジョンズホプキンス大学の研究者は「**ChatGPT Materials Explorer (CME)**」と呼ぶカスタムAIアシスタントを開発し、材料科学者向けの強力な助手として活用しています^{37 38}。CMEはChatGPTのカスタム機能で構築されたモデルで、**実際の材料データベースや物理モデルへのアクセス権**を与えられているのが特徴です^{39 40}。このツールにより、研究者は「ある組成を持つ超伝導体を設計したい」といった問いを投げかけると、従来のChatGPTでは汎用知識からの外れな回答をしがちだったところ、CMEは**専門データから根拠ある予測**を提示できます^{41 42}。実際、CMEは**巨大な材料データベースを横断検索**し、物質の結晶構造や特性を物理モデルで評価して、数秒で材料特性を予測したり関連論文を抽出したりします^{38 40}。開発者のChoudhary教授によると、ChatGPTでは最大39%程度あったハルシネーション率が、CMEでは**材料科学の正確な情報源を参照することで大幅に低減**し、専門家が「この回答は信頼できる」と思えるレベルに達したといいます⁴⁰。このように、**汎用LLMに実験データと物性計算を統合**するアプローチが材料開発にもたらす利点は計り知れません。実際の材料探索においても、AIは人間の直感では想像しにくい**斬新な化合物の組み合わせ**を提案し始めています。Google・DeepMindの研究では、AIが20000件以上の既知結晶構造データを学習し、**200万種を超える仮想化合物**を生成、その中から**38万種もの新規安定材料**を予測した例があります^{43 44}。このAIが生み出した「素材カタログ」は人間の直観に反する組み合わせも多く含み、現在ロボット実験システムで実際に合成・検証が進められています^{43 45}。わずか17日間の自動実験で、ターゲットとした58種中41種の新物質を合成する成果も報告されており、従来なら数ヶ月～数年かかる発見プロセスが**飛躍的に短縮**されています⁴⁵。専門家は「これは材料発見のChatGPTとも言える」と述べ、AIが生成した膨大な材料地図が**世界を変える可能性**に言及しています^{44 46}。さらに、こうしたAI生成の材料候補データはオープンに共有され、研究者コミュニティ全体で追試・活用が進んでいます⁴⁷。このように、最新世代のAIは材料開発において**データ駆動の高速サイクル**を実現し、これまで探索し尽くせなかった広大な化学空間に光を当て始めました。もっとも、材料科学でも**AIの予測はあくまで仮説**であり、実際に安定して有用かは実験検証が必要です^{48 49}。AIが提案したレシピ通りに合成してもうまくいかない場合もあり、その際は人間の介入で条件を微調整することでようやく目標物質が得られるケースもあります⁴⁹。今後、AIモデル自体が化学反応の動的な振る舞いまで理解を深めることで、さらに正確な予測と合成計画が立案できると期待されています。総じて、材料開発分野では**生成AIが新材料の可能性を爆発的に拡げ**、従来の試行錯誤型研究を加速・補完する存在となりつつあります。その一方で、人間研究者との協働による**結果の検証と知見の解釈**が引き続き重要な位置を占める点も共通認識となっています。

製造業における設計・試作への応用と成果



製造業のプロダクト設計や試作プロセスにも、生成AIが**デザイン革新と効率化**をもたらしています。特に注目されるのが**ジェネレーティブデザイン（Generative Design）**と呼ばれる手法で、AIが指定した制約条件のもと**無数の設計オプションを自動生成・最適化**してくれるものです⁵⁰。従来、製品設計は要求定義から構想設計・詳細設計・解析・試作・評価という段階を順に経ており、各段階で人的な反復作業が発生するため時間とコストを要していました⁵¹⁵²。ジェネレーティブデザインAIは、このプロセスを劇的に短縮します。例えばエンジニアが「軽量で強度を満たすエンジンブラケットを設計せよ」「この車両に取り付けるアームの形状を最適化せよ」といった**要件と制約を入力**すると、AIは高度なアルゴリズムで何百何千通りもの形状案を生成します⁵³⁵⁴。その上で各案をシミュレーション解析して**強度や重量といった性能評価**を自動で行い、設計案をランク付けして提示します⁵⁵⁵⁶。人間のデザイナーは上位の候補から創意工夫を加えつつ選択・改良を行えばよく、従来は**数週間かかった試作検討が数日以内に完了**する例も報告されています。実際、AltairやAutodesk、Siemensといった大手CAD/CAEベンダーは自社ソフトにジェネレーティブAI機能を搭載し始め、**自然言語でデザインの要件を伝えるだけで高度なモデリングが可能**になるツールを提供しています⁵⁰⁵⁷。この技術により、設計者は**対話的にAIから提案を引き出しつつデザイン空間を探索**できるようになりました⁵⁸⁵⁹。例えばある産業用車両のアタッチメント設計では、AIが提示した多様な形状（フォークリフト型、ブラウ型など）を人間が評価し、ベストな案を短時間で決定できたケースがあります。従来は設計者の経験則に頼っていた**勘所の探索**を、AIが漏れなく実施してくれるイメージです。さらに**リアルタイムシミュレーションやデジタルツイン**と組み合わせることで、AIが提案した設計の実製造上の課題や性能を即座に検証することも可能になっています⁶⁰⁶¹。NVIDIAの報告によれば、最新のワークステーション環境ではフォトリアルなレンダリングをリアルタイム表示しつつ**数千のデザインバリエーションを短時間で評価**でき、航空宇宙や自動車産業での導入が進んでいます⁶²⁶³。ジェネレーティブデザインは単に形状を提案するだけでなく、**製造方法（工法）自体の最適化**も支援します⁶⁴。AIは生成したデザインが実際に加工可能か（例：3Dプリントに適する形状か、部品点数を減らせないか）まで考慮し、**製造工程まで含めた包括的な提案**を行います⁶⁵。これにより、斬新なデザインが生産段階で不適切だったという無駄を減らし、設計から量産までのシームレスな流れを構築できます。GoogleのGemini 3も「**プロンプト一つで完全なフロントエンドUIをプロトタイプし、エージェント的コーディングで製品化まで一気通貫する**」デモを示すなど⁶⁶⁶⁷、ソフトウェア面でのプロトタイピングを飛躍的に加速しています。製造業ではハードとソフトの両面で、**生成AIが試作品開発のサイクルを前倒しし、設計の自由度を広げている**のです。加えてGemini 3は**複数のツールを駆使した計画立案**にも優れており、工場の生産計画やサプライチェーン最適化といったタスクも自動化しつつあります⁶⁸⁶⁹。例えば、複数センサーのログデータ解析から設備故障を予兆検知したり、部品在庫データをもとに動的に調達計画を組むといった高度なエージェント機能で、**製造プロセス全体の効率化**にも寄与し始めています⁷⁰⁶⁶。以上のように、最新世代の生成AIは**設計の自動化・高度化と試作フェーズの迅速化**を通じて製造業に革新をもたらしています。人間の設計者はクリエイティブな判断や美的評価により集中でき、反復作業はAIに任せることで**製品開発のスピードと品質が向上**しているのです。

残る課題と専門家の見解

最新世代の生成AIが多くの課題を克服しつつある一方で、**完全な解決には至っていない問題**も残存しており、専門家は慎重な姿勢を崩していません。まず**ハルシネーションの根絶**は依然として難題です。信頼性評価の新たなベンチマークによれば、GPT-5.1やAnthropic Claudeなど主要モデルはいずれも**誤情報を自信たっぷりに答えてしまう傾向**があり、Google Gemini 3 Proでさえ誤った回答の88%はシステムが自信過剰になった「幻覚」によるものだったと報告されています^{25 71}。モデルの規模拡大によって知識網羅性は向上するものの、「**わからない**」と**保留せずに推測で答えてしまう癖**は大きく改善されておらず^{25 72}、AI開発者は評価手法を工夫してこの問題に取り組んでいます（例えば間違いへのペナルティを課す「Omniscienceスコア」で慎重さを評価するなど⁷³）。また**ブラックボックス性**の問題も根深いものがあります。モデル内部の判断理由が不透明であるため、AIの出力に誤りやバイアスが潜んでいても一見すると見抜きにくいという点です。これについては可視化やチェーン・オブ・ソートの監視など技術開発が進められていますが、現状では**重要な決定は人間が責任を持ってレビューする体制**が不可欠です^{7 22}。特に医療・製造など高リスク領域では、AIを補佐役と位置づけ**最終判断は専門家が行う**という線引きが推奨されています³⁶。加えて、**倫理・法的な説明責任**の問題も指摘されています。生成AIが提案した判断に誤りがあった場合、誰が責任を負うのか、AI自体に説明能力がない中でどう説明責任を果たすかという課題です。企業導入の現場ではこの点を踏まえ、AIの出力ログを保存・検証したり、判断根拠を逐一記録する仕組みづくりが模索されています^{73 74}。例えば法律業務AIでは、生成AIが作成した契約要約について必ず元の契約文書への参照リンクを提示する設計とし、誤解や不透明な省略を防ぐ努力がなされています。**偏見（バイアス）**の問題も依然残ります。大規模モデルは学習データ由来のバイアスを内包し得るため、専門家は「AIが生成した提案が、多様性や公正性の観点で問題を含まないか**慎重にレビューすべき**」と強調しています。実際、医薬品候補提案でも商業的利害や過去の知見に偏りすぎた選択肢をAIが提示する可能性があり、そうした場合には人間がバランスを取る必要があります。総合すると、最新モデルは**驚異的な能力向上**を遂げたものの「万事解決」とはいかず、**人間との協調とガバナンス**が引き続き重要だというのが専門家の見解です。OpenAIも「モデルは依然不完全であり、クリティカルな用途では常にダブルチェックを」と明言しています⁶。今後、説明可能なAI（XAI）の技術や誤情報検出の自動化がさらに発展すれば、これら残る課題も徐々に解消されていくと期待されます。それまでは**AIの利点を最大限活かしつつ人間の責任ある判断で補う**というハイブリッドな運用が求められるでしょう。

結論

GPT-5.2やGemini 3 Deep Thinkに代表される最新世代の生成AIは、**かつて指摘されていた弱点を大きく克服しつつ、研究開発の現場に革新をもたらしています**。創薬ではAIが膨大な医学知識を駆使して新たな治療仮説を提案し、材料科学ではAIが人知を超える組み合わせで新物質候補を発見、製造業ではAIが設計プロセスを高速化・高度化するなど、その応用は多岐にわたります。従来の生成AI否定論で挙げられた誤情報生成も徐々に減少しつつあり、創造性の限界は大幅に押し広げられ、再現性や説明性も改善の兆しが見えています。しかし同時に、**万能ではないAIを人間のツールとして責任を持って使いこなす枠組みづくり**が不可欠です。最新モデルによって半歩現実に近い「汎用AI」の可能性を最大化するために、技術者・専門家・政策立案者が協調し、安全で効果的な活用指針を整備していくことが求められています。^{7 36} 革新的なAIはR&Dの飛躍を支えつつありますが、その真価を引き出すのも制御するのも人間次第であり、**新世代AIと人間の共創による未来**が今まさに始まろうとしています。

1 2 3 **OpenAI launches GPT-5.2 after 'code red' push to counter Google's Gemini 3 | Reuters**

<https://www.reuters.com/technology/openai-launches-gpt-52-ai-model-with-improved-capabilities-2025-12-11/>

4 5 6 19 **Introducing GPT-5.2 | OpenAI**

<https://openai.com/index/introducing-gpt-5-2/>

7 12 13 14 15 16 20 21 22 23 **Early experiments in accelerating science with GPT-5 | OpenAI**

<https://openai.com/index/accelerating-science-gpt-5/>

8 66 67 68 69 70 **Gemini 3 is available for enterprise | Google Cloud Blog**

<https://cloud.google.com/blog/products/ai-machine-learning/gemini-3-is-available-for-enterprise>

9 10 **Here's What Happened When We Tried Gemini 3 "Deep Think" and Google's No-Code Agents**

<https://www.marketingaiinstitute.com/blog/gemini-3-deep-think>

11 17 18 **Gemini 3 - Google DeepMind**

<https://deepmind.google/models/gemini/>

24 25 71 72 73 74 **Gemini 3 Pro tops new AI reliability benchmark, but hallucination rates remain high**

<https://the-decoder.com/gemini-3-pro-tops-new-ai-reliability-benchmark-but-hallucination-rates-remain-high/>

26 27 28 35 36 **OpenAI Launches GPT-5 with Healthcare Focus as Altman Champions Medical**

Applications

<https://hlth.com/insights/news/openai-launches-gpt-5-with-healthcare-focus-as-altman-champions-medical-applications-2025-08-08>

29 30 31 32 33 34 **A new playbook for drug repurposing - Blog | OpenAI Academy**

<https://academy.openai.com/public/blogs/using-gpt-5-pro-repurposing-existing-drugs-2025-12-11>

37 38 39 40 41 42 **ChatGPT Tool Improves Materials Properties Prediction | Technology Networks**

<https://www.technologynetworks.com/informatics/news/chatgpt-lab-assistant-predicts-material-properties-in-seconds-405027>

43 44 45 46 47 48 49 **A Google DeepMind AI Just Discovered 380,000 New Materials. This Robot Is Cooking Them Up.**

<https://singularityhub.com/2023/11/30/a-google-deepmind-ai-just-discovered-380000-new-materials-this-robot-is-cooking-them-up/>

50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 **Transforming Product Design Workflows in Manufacturing with Generative AI | NVIDIA Technical Blog**

<https://developer.nvidia.com/blog/transforming-product-design-workflows-in-manufacturing-with-generative-ai/>