

ChatGPT 4o

共有する

Hi

あなた:

NVIDIA AI Summit Japan .docx

ドキュメント

日本語で要約してください。

ChatGPT:

ChatGPT

NVIDIA AI サミット日本でのジェンスン・ファン CEO の特別講演では、NVIDIA のシミュレーション技術と AI への貢献が紹介されました。日本は NVIDIA にとって重要な市場であり、SEGA や東京工業大学との協力を通じて、初期の画期的な進展が多く生まれました。また、CUDA を使ったスーパーコンピュータや、任天堂の Switch 向けのプロセッサ開発など、多くの「初」が日本で実現されています。

ファン氏は、CPU と GPU の協調による「アクセラレーテッド・コンピューティング」の重要性を強調し、AI の大規模な進化とそれに伴う計算リソースの必要性について説明しました。また、AI エージェントのライフサイクル管理システム「NeMo」、物理 AI のための「Omniverse」、産業用ロボットのトレーニングシステム「Isaac Lab」といったソフトウェアプラットフォームも紹介され、将来の AI 産業を支える基盤として位置づけられています。

講演の中で、ソフトバンクと協力し、日本国内最大の AI ファクトリー構築計画を発表。新しい AI インフラ「AI-RAN」を構築し、ソフトバンクの基地局を通じて AI を全国に分配することが予定されています。

4o

あなた:

「日本は NVIDIA にとって重要な市場であり、SEGA や東京工業大学との協力を通じて、初期の画期的な進展が多く生まれました。また、CUDA を使ったスーパーコンピュータや、任天堂の Switch 向けのプロセッサ開発など、多くの「初」が日本で実現されています。」をくわしく

ChatGPT:

ChatGPT

NVIDIA にとって、日本は初期から非常に重要な市場でした。ジェンスン・ファン CEO の講演でも触れられたように、NVIDIA の成功には、日本企業や機関との協力が大きな役割を果たしています。ここでは、いくつかの具体的な事例について詳しく説明します。

1. SEGA との協力

SEGA は、NVIDIA がまだ小さなスタートアップ企業だった頃から、3D グラフィックス技術の可能性を認識し、協力を進めた日本企業の一つです。特に、著名なゲーム開発者である鈴木裕氏が率いるチームは、NVIDIA の GPU を使って SEGA の人気ゲームの 3D 版を開発することで、当時のゲーム業界において画期的な進展をもたらしました。これは、NVIDIA がゲーム業界での立ち位置を確立するきっかけにもなりました。

2. 東京工業大学との連携によるスーパーコンピュータ開発

東京工業大学(Tokyo Tech)は、NVIDIA が CUDA(Compute Unified Device Architecture)というプログラミングモデルを開発した際、その技術を使ってスーパーコンピュータ「TSUBAME 1.2」を構築しました。CUDA は、GPU の計算能力を活用するために開発されたソフトウェアプラットフォームで、科学技術計算の分野で特に有用です。TSUBAME 1.2 は、科学研究の進展に貢献しただけでなく、CUDA の普及と GPU を用いた並列計算の可能性を世界に示す重要な役割を果たしました。

3. 任天堂 Switch 向けプロセッサの開発

NVIDIA は、日本の大手ゲームメーカーである任天堂と協力して、Nintendo Switch 向けのカスタムプロセッサを開発しました。このプロセッサは、NVIDIA のモバイル向け GPU 技術をベースにしており、持ち運び可能なゲーム機としての Switch のパフォーマンスを支える鍵となっています。この開発は、ゲーム機市場における NVIDIA の存在感をさらに高め、日本市場における重要性を再確認する出来事でした。

これらの取り組みを通じて、日本は NVIDIA にとって単なる市場を超えて、イノベーションと新技術の実験場、さらには最初の採用者として重要な役割を果たしてきました。これにより、日本市場での実績が NVIDIA の他地域での展開や技術の進化を後押しする形となり、同社にとって欠かせないパートナーシップの一つとなっています。

4o

あなた:

「CPU と GPU の協調による「アクセラレーテッド・コンピューティング」の重要性を強調し、AI の大規模な進化とそれに伴う計算リソースの必要性について説明しました。」をくわしく

ChatGPT:

ChatGPT

NVIDIA のジェンスン・フアン CEO は、「アクセラレーテッド・コンピューティング」と呼ばれる CPU と GPU の協調動作がもたらす新しい計算モデルの重要性について強調しました。この技術は、AI や科学計算などの高度な演算を効率よく処理するための基盤であり、従来のコンピューティング手法からの大きな変革とされています。以下、その詳細をさらに掘り下げます。

1. アクセラレーテッド・コンピューティングの基本概念

従来、計算の多くは「シーケンシャル(逐次的)処理」に優れた CPU(中央処理装置)に依存していました。しかし、特に AI やシミュレーションのような膨大なデータを同時に処理する場面では、CPU の処理速度に限界がありました。ここで NVIDIA の GPU(グラフィックス処理装置)が注目されました。GPU は並列処理に優れており、膨大なデータを同時に処理することで、計算速度を大幅に向上させることができます。

アクセラレーテッド・コンピューティングでは、CPU がシーケンシャルな処理を担当し、GPU が並列処理を担当することで、各々の強みを最大限に活かしながら協調して動作します。これにより、従来の計算モデルでは数時間、数日かかっていた演算が数分、あるいは数秒で実行できるようになり、特に AI のトレーニングや推論において計算効率が劇的に向上しました。

2. AI の進化と計算リソースの増加

フアン氏は、AI モデルが進化するに伴い、必要な計算リソースが指数関数的に増加していると指摘しました。AI 分野では、「スケーリング法則」として知られる現象があり、モデルの規模を 2 倍にすると、必要なデータ量が 2 倍、そのデータの処理に必要な計算量が 4 倍に増加します。こうした法則により、AI の進化には、従来の CPU のみの計算能力では不十分であり、GPU の持つ並列処理能力を組み合わせたアクセラレーテッド・コンピューティングが必要不可欠とされています。

3. 二つのスケーリング法則

フアン氏は、AI のトレーニングと推論の両方にスケーリング法則があることも説明しました。トレーニングスケーリング法則は、AI モデルをより効果的に、より高品質にするために、ますます多くの計算リソースを必要とすることを示しています。一方で、推論スケーリング法則は、AI が「考える」プロセスを効率化するための法則であり、より良い推論結果を出すためには、推論のための計算リソースも増やす必要があるというものです。

4. アクセラレーテッド・コンピューティングの持続可能性

このように AI の進化には膨大な計算能力が求められますが、NVIDIA はこの計算効率の向上と共にエネルギー消費やコストの低減にも注力しています。新しい GPU アーキテクチャを開発するたびに、性能を向上させつつも消費電力やコストを大幅に削減することを目指しており、性能向上がエネルギー効率や経済性の改善につながるよう設計しています。これにより、AI の進化を支えつつ、持続可能な形で計算リソースを提供できるよう努めています。

このように、アクセラレーテッド・コンピューティングは、単に速度を上げるだけでなく、AI の発展に不可欠なリソースを効率よく提供するための基盤であり、今後の AI 技術の進展に大きく寄与するものとされています。

4o

あなた:

「AI エージェントのライフサイクル管理システム「NeMo」、物理 AI のための「Omniverse」、産業用ロボットのトレーニングシステム「Isaac Lab」といったソフトウェアプラットフォームも紹介され、」をくわしく

ChatGPT:

ChatGPT

NVIDIA のジェンスン・ファン CEO は、AI エージェントやロボットの開発を支えるための重要なソフトウェアプラットフォームとして、「NeMo」、「Omniverse」、そして「Isaac Lab」を紹介しました。それぞれのプラットフォームは AI エージェントとロボットのライフサイクルにおける異なる段階を支援するために設計されており、以下のような役割を果たしています。

1. NeMo: AI エージェントのライフサイクル管理システム

NeMo は、NVIDIA が開発した AI エージェントのライフサイクルを管理するプラットフォームです。AI エージェントは、マーケティング、顧客サポート、研究支援など、さまざまな業務を遂行できる「デジタル従業員」として機能します。NeMo は、これらのエージェントを以下のステップで管理・育成するためのシステムを提供します。

- **データキュレーション:** エージェントのトレーニングに必要なデータを収集・整理します。
- **トレーニングとファインチューニング:** 特定のタスクに必要なスキルを AI に教え、微調整を行います。
- **評価:** トレーニングが終了したエージェントの能力を評価し、必要に応じて修正します。
- **ガードレールの設定:** エージェントが指定されたタスクのみを実行し、不要な業務に干渉しないようにするための制御機能を提供します。

また、NeMo は、これらのプロセスを支援するためのライブラリ群を提供し、エコシステム内の他のサービスや企業と連携して、AI エージェントの導入・運用を容易にしています。

2. Omniverse: 物理 AI(ロボット)の仮想訓練プラットフォーム

Omniverse は、NVIDIA が開発したデジタルツインや仮想シミュレーションのためのプラットフォームです。物理 AI(ロボット)は、実世界での業務をこなすために高度な訓練が必要ですが、実際の環境での訓練にはコストやリスクが伴います。Omniverse は仮想環境を提供し、ロボットが安全かつ効率的に学習・訓練を行えるようにします。

Omniverse は、以下のような機能を備えています。

- **デジタルツインの構築:** 実際の物理空間の詳細なモデルを作成し、ロボットが仮想空間で試行錯誤を行いながら学習できる環境を提供します。
- **合成データ生成:** 現実のデータが不足している場合でも、仮想環境内でさまざまなシナリオを生成し、ロボットの学習データとして利用します。
- **物理ベースのシミュレーション:** ロボットがリアルな物理法則に基づいて動作するため、実際の運用に近い訓練が可能です。

このプラットフォームを使うことで、企業は物理 AI の訓練にかかるコストを大幅に削減できるだけでなく、効率的にロボットのスキルを向上させることができます。

3. Isaac Lab: 産業用ロボットのトレーニングシステム

Isaac Lab は、産業用ロボットのための訓練プラットフォームで、Omniverse 上で稼働するロボット向けの「トレーニングジム」ともいえる存在です。Isaac Lab は、特にロボットが複雑な作業を遂行するために必要なスキルを習得するのに役立つ機能を備えています。Isaac Lab には、次の 3 つの主要なワークフローが含まれます。

- **Groot-Mimic:** 人間の動作をロボットが模倣するためのシステムで、人間のデモンストレーションに基づいてロボットが同様の動きを習得し、さらに多様な場面に応用できるようにします。
- **Groot-Gen:** ジェネレーティブ AI 技術を用いて、ロボットがさまざまな環境や条件下で行動を試みるためのシナリオを生成します。これにより、ロボットは多様な状況に適応する能力を向上させることができます。
- **Groot-Control:** ロボットが学習したスキルを 1 つの統合モデルに蒸留し、動作スキルを効率よく管理するためのフレームワークです。

これにより、ロボットは現場で自律的に業務を遂行できるようになり、未来の工場ではロボットが主体となり作業を進めることが可能となります。

このように、NeMo、Omniverse、Isaac Lab はそれぞれが AI エージェントや物理 AI の開発・運用において重要な役割を果たしており、NVIDIA はこれらのプラットフォームを通じて、あらゆる業界で AI 技術の導入と応用を支援しています。

40

あなた:

「講演の中で、ソフトバンクと協力し、日本国内最大の AI ファクトリー構築計画を発表。新しい AI インフラ「AI-RAN」を構築し、ソフトバンクの基地局を通じて AI を全国に分配することが予定されています。」をくわしく。

ChatGPT:

ChatGPT

NVIDIA とソフトバンクは、日本国内で最大規模の AI ファクトリーを構築する計画を発表しました。この取り組みは、AI と無線アクセスネットワーク(RAN)を統合する新しいインフラストラクチャ「AI-RAN」の構築を中心としています。AI-RAN は、AI アプリケーションと無線アクセスネットワークを同一のコンピューティング基盤上で統合する新しいアーキテクチャであり、これにより、ソフトバンクの基地局を通じて AI サービスを全国に提供することが可能となります。

[ASCII.jp](#)

この計画の一環として、ソフトバンクは NVIDIA の最新チップ「Blackwell」を搭載した国内最大級の AI スーパーコンピューターを構築し、さらに「Grace Blackwell」チップの導入も計画しています。これらのハードウェアは、AI-RAN の基盤となり、高性能でエネルギー効率に優れたソフトウェア定義の RAN を提供します。

[ロイター](#)

AI-RAN の導入により、ソフトバンクは 5G ネットワークと AI サービスを統合し、通信事業者が新たな収益機会を得ることを目指しています。具体的には、製造ロボットや自動運転車の遠隔操作、製造業や農業におけるコンピュータービジョン、生成 AI 駆動型のコパイロットやパーソナルアシスタントなど、多様な AI アプリケーションの提供が可能となります。

[NVIDIA ブログ](#)

さらに、ソフトバンクは国内で安全な AI コンピューティングの需要に応えるため、NVIDIA AI Enterprise を活用し、AI マーケットプレイスの構築も目指しています。これにより、地方自治体や大企業、スタートアップ企業が、セキュリティ、パフォーマンス、現地の言語や文化基準に合わせた生成 AI アプリケーションを構築、開発、展開することが可能となります。

[ASCII.jp](https://www.asci.jp)

このように、NVIDIA とソフトバンクの協力は、日本における AI 産業の発展と通信インフラの革新を推進し、AI 時代における新たなビジネスチャンスの創出に寄与することが期待されています。

情報源



4o

あなた:

ロボットのところで、日本企業とのかかわりを話していました。そのあたりをくわしく。

ChatGPT:

ChatGPT

NVIDIA は日本企業と協力し、ロボティクス分野での革新を推進しています。特に、安川電機やソフトバンクとの連携が注目されています。

安川電機との協力: 安川電機は、NVIDIA のテクノロジーを活用して AI 対応ロボットの開発を進めています。2023 年 11 月に東京で開催された国際ロボット展 (iREX) では、NVIDIA と安川電機を含む複数のパートナーが最先端のロボティクスソリューションを展示しました。この展示では、NVIDIA の Isaac Nova Orin アーキテクチャをベースとしたリファレンスデザインロボット「Nova Carter」が日本で初公開され、産業用ロボットの未来像が示されました。

[NVIDIA ブログ](#)

ソフトバンクとの協力: ソフトバンクは、NVIDIA の最新チップ「Blackwell」を搭載した AI スーパーコンピューターを導入し、AI と 5G ネットワークを統合する新しいインフラ「AI-RAN」の構築を計画しています。この取り組みは、製造ロボットや自動運転車の遠隔

操作、製造業や農業におけるコンピュータービジョン、生成 AI 駆動型のコパイロットやパーソナルアシスタントなど、多様な AI アプリケーションの提供を目指しています。

ロイター

これらの協力関係は、日本のロボティクス産業の発展と AI 技術の普及に大きく寄与しています。

情報源



ChatGPT 4o

共有する

Hi

あなた:

NVIDIA AI Summit Japan .docx

ドキュメント

ステップバイステップで要約を作成してください。要約に当たっては、重要な言葉は出来るだけ本人の口調を活かすようにしてください。

ChatGPT:

ChatGPT

以下は「NVIDIA AI サミット ジャパン」でのジェンスン・ファン CEO の特別講演内容のステップバイステップの要約です。彼の口調をできるだけ活かして要約しています。

1. シミュレーション技術の力:

- NVIDIA は「物理、仮想世界、知能」をシミュレーションするテクノロジー企業であり、このシミュレーションにより未来を予測する「タイムマシン」を構築しています。
- このサミットは、日本のエコシステムとの連携の重要性を強調する場であり、350 社以上のスタートアップや数多くの企業が参加しています。

2. 日本と NVIDIA の歴史:

- NVIDIA は日本で数多くの「初」を経験。例えば、最初のゲーム開発パートナーは日本の SEGA であり、CUDA を使用した最初のスーパーコンピュータ「つばめ 1.2」も東京工業大学で実現。
- モバイルプロセッサ技術も日本で初めて導入され、後に任天堂スイッチのプロジェクトに貢献。

3. 新しい計算モデル「アクセラレーテッドコンピューティング」:

- GPU と CPU が協力する新しい計算モデルが「アクセラレーテッドコンピューティング」であり、従来のシーケンシャルな計算だけではなく、並列処理が可能になりました。
- これにより、膨大な計算負荷をこなすためのライブラリやアルゴリズムを開発し、AI や科学計算の発展に寄与しています。

4. **AI ライブラリと新しいアルゴリズム:**
 - NVIDIA の多様な AI ライブラリ (CuLitho、CuDSS、CuOPT、CuDNN など) が紹介され、特に CuDNN がディープラーニングの普及を支えたことで AI の発展を 100 万倍スケールアップすることができました。
5. **ソフトウェアの進化と「ソフトウェア 2.0」:**
 - ソフトウェア開発は、従来の人間がコードを書く「ソフトウェア 1.0」から、機械学習を活用してコンピュータが学習する「ソフトウェア 2.0」へと進化しました。
 - これにより、AI は膨大なデータから予測や推論が可能となり、ニューラルネットワークが現代の OS として機能しています。
6. **多様な AI エージェントと産業の革命:**
 - 「AI エージェント」と「物理 AI (ロボティクス)」が 2 つの主要な AI 応用分野であり、AI エージェントはマーケティング、カスタマーサポート、サプライチェーンなどで活躍します。
 - 日本のロボティクス技術 (川崎重工、ファナックなど) と AI の融合により、ロボティクス革命を先導する可能性があります。
7. **NVIDIA とソフトバンクの AI インフラ協力:**
 - ソフトバンクと協力して、日本国内最大の AI ファクトリーを建設し、25 エクサフロップスの計算能力を持つ予定。
 - さらに、AI-RAN を介して AI を広範囲に分配するネットワークを構築し、AI の社会的・経済的な影響を最大化する計画が発表されました。
8. **AI の新しいインフラストラクチャ:**
 - AI は新しい「インフラ」として機能し、日本全土に AI グリッドが構築され、研究者、スタートアップ、企業が新たなサービスや産業を創出できる基盤が整備されます。
 - 道路や電力のように AI が次世代の産業革命を支える重要なインフラとして位置づけられています。
9. **将来展望と AI の国家戦略:**
 - AI の国家レベルでの活用が不可欠であり、各国は AI データセンターを構築し、自国のデータを利用して AI を生み出す必要があると強調。
 - 日本における AI 技術の拡充と次世代への展望が語られ、講演は日本に向けた感謝と連携のメッセージで締めくくられました。

この要約はサミットでの主な発表内容や、日本の役割に対する強調点をステップごとにまとめたものであり、ジェンスン・フアン氏の視点を反映しています。

あなた:

シミュレーション技術の力:をくわしく。要約に当たっては、重要な言葉は出来るだけ本人の口調を活かすようにしてください。

ChatGPT:

ChatGPT

シミュレーション技術の力(詳細):

ジェンスン・ファン氏が強調したのは、NVIDIA の核となる技術である「シミュレーション技術」が、どれほど多様な分野において未来を切り開く「タイムマシン」として機能するかという点です。彼はこう語ります。「NVIDIA はシミュレーション技術の会社です。私たちは物理を、仮想世界を、そして知能をシミュレーションします」。ここで、NVIDIA が行っているのは単なるアニメーションや単純なコンピュータグラフィックスではなく、「すべてがシミュレーション」ということです。

シミュレーションは、未来を予測するための強力なツールであり、例えば AI やデジタルツインの分野においては、シミュレーション技術がさまざまな産業革命を牽引する基盤となることが期待されています。彼は、「NVIDIA はタイムマシンを構築しているようなものです」と述べ、同社の技術が人々や企業に未来を予測し、その先を見据えた意思決定をサポートする役割を持つことを示唆しました。

また、このシミュレーション技術は日本との歴史とも密接に関連しており、日本市場との長い関係が、数多くの技術的「初」を生み出してきたことを強調しています。

4o

あなた:

日本と NVIDIA の歴史:をくわしく。要約に当たっては、重要な言葉は出来るだけ本人の口調を活かすようにしてください。

ChatGPT:

ChatGPT

日本と NVIDIA の歴史(詳細):

ジェンスン・ファン氏は、「日本は NVIDIA にとって特別な場所だ」と繰り返し強調しています。NVIDIA の歴史の中で、日本は数多くの「初」をともに実現してきた重要なパートナーであり、この関係がいかに深く長いものかを彼は語ります。「NVIDIA が初めてゲーム開発でパートナーを組んだのは日本の SEGA でした」。有名なゲーム開発者・鈴木裕氏と一緒に、3D ゲームを NVIDIA の GPU にポートし、3D グラフィックスの新たな可能性を開いたことを紹介しました。

また、彼は日本が「スーパーコンピュータの世界でも最初の地」だったことに触れました。東京工業大学の「つばめ 1.2」を例に挙げ、「つばめ 1.2」は NVIDIA CUDA を使用した最初のスーパーコンピュータであり、日本が科学計算を進化させる場として最先端の役割を果たしていたと述べています。さらに、モバイルプロセッサの分野でも、NVIDIA が初めて成功を収めたのは日本であり、それが「非常に愛着のあるプロジェクトである任天堂スイッチ」に結びついたことも強調しました。

こうした「日本での多くの初めて」を通じて、NVIDIA はシミュレーションと AI 技術を大きく前進させ、今の成長の礎を築きました。そして現在、日本が AI 革命の新時代に向かっていくパートナーとしてますます重要な役割を果たすと期待しています。「NVIDIA にとって日本は、創業当初から大切な存在でした」と彼は語り、長年の協力関係に感謝を表明しています。

40

あなた:

新しい計算モデル「アクセラレーテッドコンピューティング」:をくわしく。要約に当たっては、重要な言葉は出来るだけ本人の口調を活かすようにしてください。

ChatGPT:

ChatGPT

新しい計算モデル「アクセラレーテッドコンピューティング」(詳細):

ジェンスン・ファン氏が語る「アクセラレーテッドコンピューティング」は、NVIDIA の革新的なコンピューティングモデルです。このモデルの核は、CPU と GPU が協力して計算処理を行うというアプローチです。彼は、「NVIDIA は CPU を置き換えようとするのではなく、CPU を補完し、CPU が得意とするシーケンシャルな処理と、GPU が得意とする並列処理を組み合わせることを目指してきた」と述べています。

彼が特に強調するのは、この「アクセラレーテッドコンピューティング」が単なる「並列コンピューティング」とは異なる点です。「ただの並列処理ではありません。CPU と GPU が協調して動作する、全く新しい計算モデルです」。従来の CPU だけを使うコンピューティングが 1964 年から続いてきた中で、NVIDIA が提唱するアクセラレーテッドコンピューティングは根本的な変化をもたらしました。

しかし、彼はこの新しいモデルを実現するために、単に従来の CPU 向けのソフトウェアを GPU で動かすだけでは不十分であることを指摘します。そこで、NVIDIA はさまざまな用途に応じた「ドメイン特化型のライブラリ」を開発してきました。例えば、グラフィックスの世界では OpenGL がアクセラレーションを可能にしたのと同様に、AI や科学技術、製造業などの分野でも専用のライブラリを設計し、GPU の性能を最大限に引き出しています。

このアクセラレーテッドコンピューティングは、膨大な計算量を必要とする分野において大幅な高速化を実現し、科学計算やデータ解析、AI などの領域で「未来を創造する」ための強力な基盤となっています。

4o

あなた:

AI ライブラリと新しいアルゴリズム:をくわしく

ChatGPT:

ChatGPT

AI ライブラリと新しいアルゴリズム(詳細):

ジェンスン・ファン氏は、NVIDIA が開発した数多くの AI ライブラリと、それらがどれだけ産業を変革する潜在力を持つかを熱く語っています。彼は、「NVIDIA には 350 種類以上のライブラリがあり、それぞれが非常に重要な役割を果たしている」とし、それらが AI や計算分野での進歩を大幅に加速させていることを強調します。

代表的なライブラリとしてまず挙げられるのが **CuLitho** です。これは「計算リソグラフィ」に特化しており、半導体のチップの製造過程におけるマスクの作成にかかる時間を「数週間から数時間に短縮する」力を持ちます。この技術により、半導体業界では、

より複雑なリソグラフィーのアルゴリズムが可能となり、2 ナノメートルや1 ナノメートル以下の半導体技術の発展を加速させています。

次に、CuDNN (CUDA Deep Neural Network library) が紹介されました。ファン氏は「CuDNN はディープラーニングの拡大と民主化を実現した非常に重要なライブラリ」とし、AI のスケールを「100 万倍に拡大」できたのは CuDNN のおかげだと語ります。このライブラリがディープラーニングの普及を促進し、ChatGPT のような革新的な AI 技術の実現を可能にしました。

さらに、AI の「スパースソルバー」を可能にする CuDSS や、Parabricks (遺伝子解析用)、CuVS (ベクトルデータベース用)、CuPyNumeric (数値処理用)、CuDF (データフレーム処理用)、CuOPT (旅行セールスマン問題解決用)、Modulus (AI 物理シミュレーション用) など、多岐にわたるライブラリが紹介されています。各ライブラリが特化した計算や処理を可能にし、例えば CuOPT では旅行セールスマン問題のような「非常に複雑な最適化問題を何百倍も速く」解決できると説明しています。

彼はまた、Modulus のような AI 物理シミュレーションライブラリも強調し、AI を活用した物理シミュレーションがさまざまな産業において革新的な可能性を広げていると述べています。さらに、AI による大規模な数値処理やデータ分析を実現するために、NumPy や SQL のような標準ライブラリも NVIDIA の GPU 上でフルアクセラレーションされており、膨大な開発者とユーザーに利用されていることも紹介しています。

これらのライブラリ群は、NVIDIA の「CUDA エコシステム」として一体となり、AI の開発や導入を加速するための基盤として機能しています。「NVIDIA が創り出したライブラリとソフトウェアスタックがなければ、AI スーパーコンピュータの運用は不可能だった」と彼は述べ、これらのライブラリが AI や機械学習の革新を支える重要なツールであることを力説しています。

40

あなた:

ソフトウェアの進化と「ソフトウェア 2.0」:をくわしく

ChatGPT:

ChatGPT

ソフトウェアの進化と「ソフトウェア 2.0」(詳細):

ジェンスン・ファン氏が語る「ソフトウェア 2.0」は、従来のプログラムによるコーディングに依存していた「ソフトウェア 1.0」から、AIによって自己学習しながら機能を達成する新しいパラダイムへの転換を意味しています。彼は「ソフトウェア 1.0 では、プログラマーがアルゴリズムを記述して、そのコードが入力を受けて出力を予測する仕組みで、CPU 上で動作することが前提だった」と説明します。

一方、ソフトウェア 2.0 では、「コンピュータの性能が飛躍的に向上したことで、大量のデータから学習し、自ら関数を見つけ出すことができる」と述べています。これは、単なるコードではなく、ニューラルネットワークを通じてコンピュータが自律的に学習し、予測を行う新しい形式の「ソフトウェア」です。ソフトウェア 2.0 では、「コードを書かずに、膨大な観測データ(例示データ)を通じて機械学習モデルをトレーニング」することができ、これにより「ニューラルネットワークが GPU 上で動作し、驚異的なスケールで成長している」と説明しています。

この進化がもたらしたものの一つが、今日の大規模な言語モデル(Large Language Models, LLMs)であり、ファン氏はこれを「現代コンピュータの新たなオペレーティングシステム」と称しています。彼は、LLM のようなモデルが、テキストや音声、画像、動画などのあらゆるデジタルデータを「多様なモダリティ間で理解し、接続できる」と説明しています。これは、例えばテキストから画像生成、画像認識、音声の理解、さらには多言語間の翻訳といった複数の用途で応用可能である点で、これまでのソフトウェアとは異なるスケラブルで柔軟な特性を持っています。

ファン氏はまた、「ソフトウェア 2.0 は、膨大なデータからパターンや関係性を見つけ出すことで、単語の意味や文脈を理解することができる」と述べており、このアプローチがいかにして AI に対する根本的な理解を変えたかを強調しています。これにより、ソフトウェアが「データを理解し、さまざまな形式のインテリジェンスを生成」する能力が可能となり、企業や産業界においても、新たな活用の幅を広げています。

特に、彼はソフトウェア 2.0 の力を次のように表現します。「従来のソフトウェアはツールとして使用されてきたが、ソフトウェア 2.0 はタスクを実行し、実際の仕事をこなすことができる」。これにより、従来の IT 産業だけでなく、あらゆる業界が変革の波に直面していると述べ、「この新しいソフトウェアは、ただの IT の進化ではなく、すべての産業がシフトする時代が到来している」と締めくくっています。

「ソフトウェア 2.0」は、まさに NVIDIA の GPU とアクセラレーテッドコンピューティングの力を活用し、より複雑で高度な知能を持つ AI が実現可能となる未来を示すものであり、ジェンスン・ファン氏はこれを一つの「技術革命」として捉えています。

4o

あなた:

『多様な AI エージェントと産業の革命: 「AI エージェント」と「物理 AI(ロボティクス)」が 2 つの主要な AI 応用分野であり、AI エージェントはマーケティング、カスタマーサポート、サプライチェーンなどで活躍します。日本のロボティクス技術(川崎重工、ファナックなど)と AI の融合により、ロボティクス革命を先導する可能性があります。』をくわしく

ChatGPT:

ChatGPT

多様な AI エージェントと産業の革命(詳細):

ジェンスン・ファン氏は、AI 技術が新たな産業革命をもたらすとし、その中心には「AI エージェント」と「物理 AI(ロボティクス)」という 2 つの主要な AI 応用分野があると述べています。これらの AI 応用分野がもたらす革新について、彼は具体例を挙げながら詳しく語っています。

まず、「AI エージェント」は、デジタル空間における「デジタル従業員」のようなものであり、さまざまな業務や役割に適用されます。彼は「デジタル AI エージェントは理解し、計画し、そして行動することができる」とし、マーケティングキャンペーンの実施、顧客サポート、サプライチェーンの最適化、製品設計支援、さらには製薬業界での研究補助といった、幅広い分野で業務効率を飛躍的に向上させるとしています。具体的な役割に応じて訓練され、評価されることで、企業にとって有益な「デジタル従業員」としての機能を果たします。

この「AI エージェント」の利用をさらに容易にするため、NVIDIA は「NeMo(ニーモ)」「NIMs」「Blueprints(ブループリント)」といったライフサイクルプラットフォームを提供しています。NeMo は AI エージェントのデータキュレーション、訓練、評価、ガードレール設定など、エージェントの全段階で活用できるライブラリ群であり、Blueprints は AI の参照ワークフローとして、企業が迅速に AI エージェントを構築・導入できるよう支援

します。こうしたツールにより、「NVIDIA はサービスを提供するのではなく、AI エコシステムを支援し、企業が自らの AI エージェントを創り出すことを可能にする」と述べています。

次に、「物理 AI(ロボティクス)」は、AI 技術を物理的な機械に組み込むもので、彼は特に日本のロボティクス産業との連携を強調しました。日本のロボット技術は世界トップレベルであり、川崎重工やファナック、ヤスカワ電機、三菱電機などの日本企業が、世界の産業用ロボットの約 50%を製造しています。しかし、彼は「従来のロボットは非常に特化されており、用途が限られている」と指摘し、柔軟で自己学習できる AI の導入によって、より多様な作業環境に適応できるロボティクスの実現が重要であると述べています。

この物理 AI を可能にするために、NVIDIA は「3 つのコンピュータ」を構築しています。まず「AI トレーニング用コンピュータ」で AI を訓練し、次に「Omniverse」という仮想環境で AI をシミュレーションし、最後に「Jetson Thor」という実際のロボティクスプロセッサに AI を導入します。Omniverse は「デジタルツイン」としてロボットが仮想空間で訓練を積み重ねる場であり、「AI が自己学習し、より優れたパフォーマンスを発揮できるようになる」と説明しています。また、Groot-Gen、Groot-Mimic、Groot-Control といったフレームワークが Omniverse 上でロボットの動作を学習・一般化し、複雑なタスクを遂行できるよう支援します。

最後に、ジェンソン氏は「日本はロボット愛好国であり、AI とロボティクス革命のリーダーとなる絶好の立場にある」と述べ、日本の高度なメカトロニクス技術が、NVIDIA の最新 AI 技術と融合することで新たなロボティクス革命を先導する可能性を強調しました。特に、人間の動作や環境に適応できるヒューマノイドロボットや次世代の工場「ロボティックファクトリー」の実現が期待され、AI と日本のロボット技術の融合により、幅広い産業で効率と生産性の向上が進むことを彼は願っています。

こうして、AI エージェントと物理 AI が、デジタル空間と物理空間の両方で、革新的な産業変革を推進する可能性を持つことを彼は熱く語り、日本がその未来を牽引する役割を果たすことへの期待を表明しています。

40

あなた:

NVIDIA とソフトバンクの AI インフラ協力をくわしく

ChatGPT:

ChatGPT

NVIDIA とソフトバンクは、AI インフラストラクチャの分野で緊密な協力関係を築いています。この協力は、生成 AI の開発や 5G/6G ネットワークの次世代プラットフォーム構築を中心に展開されています。

生成 AI 開発向け計算基盤の構築

ソフトバンクは、国内最大級の生成 AI 開発向け計算基盤の稼働を開始しました。この基盤は、NVIDIA の Tensor コア GPU を 2,000 基以上搭載した AI スーパーコンピューター「NVIDIA DGX SuperPOD™」をベースにしています。これにより、国内最大級の計算能力を持つクラスターが実現されました。

[ソフトバンク](#)

国産大規模言語モデル(LLM)の開発

ソフトバンクの子会社である SB Intuitions 株式会社は、この計算基盤を活用して、日本語に特化した国産 LLM の開発を本格的に開始しました。2024 年内に 3,500 億パラメーターの国産 LLM の構築を目指しています。

[ソフトバンク](#)

次世代データセンターの構築

NVIDIA とソフトバンクは、生成 AI と 5G/6G 向けの次世代データセンターの構築に向けて協業しています。このデータセンターは、NVIDIA の Grace Hopper Superchip を活用し、AI と 5G の両方のワークロードを処理することを目指しています。これにより、低遅延でピーク時の運用を実現し、全体のエネルギーコストを大幅に削減することが期待されています。

[NVIDIA](#)

AI 計算基盤の拡張

さらに、ソフトバンクは約 4,000 基の「NVIDIA Hopper GPU」の整備を完了し、AI 計算基盤全体の GPU 数を約 6,000 基としました。これにより、計算処理能力の合計は 4.7

エクサフロップスとなりました。今後も GPU の増強を行い、合計 25.7 エクサフロップスの計算能力の実現を目指しています。

[ソフトバンク](#)

これらの取り組みを通じて、NVIDIA とソフトバンクは、日本国内における AI インフラの強化と、次世代の通信ネットワークの構築を推進しています。

情報源



40

あなた:

AI の新しいインフラストラクチャ:をくわしく

ChatGPT:

ChatGPT

AI の新しいインフラストラクチャ(詳細):

ジェンスン・ファン氏が強調する「AI の新しいインフラストラクチャ」は、これまでのインフラ基盤とは異なる、まったく新しい形態のインフラストラクチャです。彼は、AI インフラを「道路や電力、通信のような従来のインフラと同じく、次世代産業を支える重要な基盤」として位置付けています。この AI インフラが整備されることで、AI が産業や社会全体に浸透し、新たな経済活動と産業の発展が可能になります。

1. AI ファクトリー

- AI インフラの中心的な役割を果たすのが「AI ファクトリー」です。これは、膨大な AI モデルを生成し、育成する「AI の製造工場」として機能します。従来のソフトウェアは完成品として提供されますが、AI ファクトリーでは「AI モデルが絶えず進化し、学習し続ける」ことが可能です。彼は、AI モデルの生成が今や「24 時間 365 日稼働する新たな製造業」となり、「AI は製造され、常に更新されるもの」だと述べています。

2. AI-RAN (AI Radio Access Network)

- 次に、AI を社会に分配するための「AI-RAN」という新しいネットワークが構築されます。これは、従来の通信ネットワークに AI を組み込み、5G/6G 通信を超えて「AI データ」を配信する新しい通信インフラです。AI-RAN は、ソフトバンクの 200,000 以上の基地局を通じて、AI の処理能力を広範囲に展開し、あらゆるデバイスやユーザーに対して、AI の力をリアルタイムで提供することを可能にします。

3. AI ストア

- NVIDIA とソフトバンクは、AI の普及をさらに加速するため、「AI ストア」の構築も計画しています。この AI ストアでは、ソフトバンクやその他のパートナーが開発した AI アプリケーションやエージェントが提供され、企業や個人が AI ソリューションに簡単にアクセスできるようになります。これは「AI のアプリケーション・マーケットプレイス」のようなものであり、企業が自身の業務やサービスに応じた AI を柔軟に選択し、導入できる環境を提供します。

4. AI グリッド

- ファン氏はこれらの AI インフラが統合されることで、「日本全体に AI グリッドを構築する」ビジョンを掲げています。AI グリッドは、従来のエネルギーグリッドや交通インフラのように、産業や社会を支える基盤であり、AI の力を国全体に行き渡らせるためのインフラです。これにより、新たな経済的な可能性や雇用機会が生まれ、AI を基盤とした新しい産業革命が推進されると彼は述べています。

5. AI によるインフラのリセット

- 最後に、彼はこの AI インフラの構築が「技術のリセット」をもたらすと語りまます。従来のインフラが産業革命や IT 革命を支えたのと同様に、AI インフラが AI 時代の新たな基盤となり、「AI が情報を生み出し、処理し、提供する新しいエコシステムを可能にする」としています。このリセットによって、日本や世界各国が次世代産業の発展において競争力を高めることが期待されています。

この AI インフラの構築により、NVIDIA とソフトバンクは AI の普及と利活用を加速し、AI 技術が生活やビジネスのあらゆる場面で不可欠な要素となる未来を切り開くことを目指しています。

40

あなた:

将来展望と AI の国家戦略:をくわしく

ChatGPT:

ChatGPT

将来展望と AI の国家戦略(詳細):

ジェンスン・ファン氏が語る「将来展望と AI の国家戦略」は、AI が国や産業の基盤となる時代を見据えたビジョンです。彼は、AI 技術が単なる技術革新ではなく、各国にとって「国家の繁栄と安全保障を支える新たな柱」になることを強調しています。特に日本のような国が AI 時代にどのように対応すべきか、彼の期待や洞察が語られました。

1. AI の自国生産の必要性

- ファン氏は、各国が AI 技術を「自国で生産・管理する重要性」を強調しています。AI は単なるソフトウェアやハードウェアの枠を超えた存在であり、国民のデータや知識、文化と深く結びついています。そのため、AI モデルやデータセンターを「国家資源」として管理し、国内で AI を製造・維持することが求められるとしています。
- 「自国のデータを処理し、自国民のための AI を生み出すことが必要です。他国にこれを委託するのは、国の知識や文化の資産を他国に預けるようなもの」と述べ、AI が国家戦略の一環として保護されるべき存在であることを強調しました。

2. インフラ整備と国家の競争力

- ファン氏は「AI インフラがなければ、AI 産業の成長は難しい」と述べ、国家レベルでのインフラ整備の必要性を強調しています。彼は「自動車産業には道路が不可欠のように、AI 産業には AI インフラが必要だ」と例え、各国が競争力を高めるためには、強力な AI インフラ(AI ファクトリー、AI-RAN、AI グリッドなど)を持つことが不可欠であると述べています。

- 特に、日本が AI 時代のリーダーシップを確立するためには、AI ファクトリーや AI グリッドの整備によって、AI 技術が経済全体に影響を与える基盤を確立することが重要です。

3. 教育・研究機関の支援

- AI の国家戦略には、若い世代や新興企業、研究者への支援も含まれるべきだとし、ファン氏は「AI インフラを大学やスタートアップが利用できるようにし、次世代のイノベーションを促進する必要がある」と述べています。彼は、NVIDIA とソフトバンクが日本国内の教育機関やスタートアップに対して AI データセンターの利用を提供し、優れた計算資源を低コストで活用できるプログラムを提供することに期待しています。
- これにより、日本の AI 研究や開発のスピードを加速させ、国内での AI 人材育成と革新を促進する基盤が整えられるとしています。

4. AI と安全保障

- ファン氏は AI が国防や安全保障にも大きな影響を与えると指摘し、「各国が独自の AI インフラを持ち、国家データを安全に保護することは、国家安全保障の観点からも重要」と述べています。AI によるデータ処理や意思決定の自動化が進む中で、AI 技術が軍事や防衛の分野にも活用される可能性があり、各国がその管理体制を強化する必要があると考えられています。

5. 日本の役割と可能性

- 最後に、ファン氏は日本が AI 革命のリーダーとして大きな潜在力を持つと述べています。日本は、過去に製造業やメカトロニクスで世界をリードしてきた技術大国であり、特にロボティクスや半導体技術において高い競争力を持っています。ファン氏は、「日本はリセットの機会を活用して、AI 技術においても世界をリードする可能性がある」と述べ、今が「新たな産業革命におけるリーダーシップを確立する好機」であるとしています。
- ファン氏は、日本が国家戦略として AI インフラを整備し、AI 時代の「先頭に立つ」ことを期待しており、NVIDIA とソフトバンクのパートナーシップがその実現に大きく貢献することを確信しています。

結論

このように、ファン氏のビジョンは、AI が国家の競争力と安全保障を左右する時代の到来を見据えています。彼は、各国が AI の持つ潜在力を最大限に引き出し、自国民のために活用するための体制を整えるべきだとし、日本がそのモデルケースとなることを期待しています。

4o

あなた:

講演のなかで、ソフトバンクの孫会長が登場し、対談になっています。その内容をステップバイステップでくわしく、できるだけ本人の言葉を用いて、要約してください。

ChatGPT:

ChatGPT

2024 年 11 月 13 日、東京で開催された NVIDIA AI サミットにおいて、NVIDIA の創業者兼 CEO であるジェンソン・ファン氏と、ソフトバンクグループの代表取締役会長兼社長執行役員である孫正義氏が対談を行いました。この対談では、AI の未来、生成 AI の可能性、日本の技術的自立、そして両社の協力関係について深く議論されました。

1. AI の未来と生成 AI の可能性

孫氏は、AI が人類の知能を超える「人工超知能 (ASI)」の実現を目指していると述べました。彼は、「AI は人類の知能を超え、1 万倍の知能を持つ ASI の実現を目指しています」と語り、AI が医療、交通、教育など多岐にわたる分野で革命をもたらす可能性を強調しました。ファン氏もこれに同意し、「生成 AI は、私たちの生活や産業を根本的に変える力を持っています」と述べ、AI の進化がもたらす変革について期待を示しました。

2. 日本の技術的自立とソブリン AI

対談では、日本が AI 分野で技術的自立を達成する重要性についても議論されました。孫氏は、「日本が独自の AI インフラを持つことは、国家の競争力と安全保障に直結します」と述べ、ソブリン AI (国家主権 AI) の構築が必要であると強調しました。ファン氏もこれに賛同し、「日本が AI のリーダーシップを取るためには、自国で AI を開発し、運用する能力が不可欠です」と述べました。

3. NVIDIA とソフトバンクの協力関係

両氏は、NVIDIA とソフトバンクの協力関係についても言及しました。孫氏は、「私たちは NVIDIA と共に、AI と 5G/6G を融合させた次世代のデータセンターを構築しています」と述べ、両社の協力が日本の AI インフラの強化に寄与していることを強調しました。ファン氏も、「ソフトバンクとのパートナーシップは、AI の普及と革新を加速させる重要な要素です」と述べ、今後の協力を期待を示しました。

4. 過去の投資と今後の展望

孫氏は、過去に NVIDIA への投資を行い、その後売却したことについて、「NVIDIA の成長を見て、売却したことを後悔しています」と述べました。しかし、彼は今後の AI 分野での投資と協力に対する意欲を示し、「これからも NVIDIA と共に、AI の未来を切り開いていきたい」と語りました。ファン氏も、「孫さんとの協力は、AI の可能性を最大限に引き出すための鍵となります」と述べ、両社の連携がもたらす未来への期待を表明しました。

この対談を通じて、NVIDIA とソフトバンクが AI 分野での協力を深化させ、日本の技術的自立と AI の未来に向けた取り組みを強化していく姿勢が明確に示されました。

情報源