

人工知能 (AI)

パートナー支援パッケージ

インテル製品搭載ソリューションで顧客のビジネス課題に対応する

AI をあらゆる場所に導入

あらゆるプラットフォームでAIコンティニュームを実現
クライアントからエッジ、データセンター、クラウドまで

インテルのパートナーになる理由

インテルには、地球上のあらゆる人の生活とあらゆる企業の成果を向上させる機会がありますが

それは一社のみでは実現しません。

インテルはパートナーと共に、**AI をあらゆる場所に導入**し、導入のリスクを最小限に抑えることで、顧客に真の価値を創造します



インテルのパートナーになることで、AI エコシステム全体とパートナーに

インテルの AI 実現テクノロジーの幅広いポートフォリオと、ハードウェア、ソフトウェア、システム・インテグレーターが緊密に連携するパートナーシップにより、業界、企業、コミュニティに差別化されたビジネスの成果をもたらす、実環境のソリューションを創造します。

そしてお客様のビジネスの成長も支援します。

AI をあらゆる場所に導入する取り組みにぜひ参加してください

インテル® AI の業界への影響



「インテルは、数百億ドル規模の TAM (Total Addressable Market: 獲得可能な最大市場規模) につながる千載一遇のビジネスチャンスを理解し、あらゆる業界とビジネスセグメントに広がる AI に必要なインフラストラクチャーの構築に奔走してきました」



「業界が信じるなら、2024年は AI PC の年となります。そのすべてはインテルから始まるでしょう」



「パフォーマンスが強化された AI PC と新しいデータセンター CPU を同時にリリースしたことで、インテルが AI における独自の優位性を徐々に固めていると認識し始めています」



「重要なのは、インテルの新チップが予定通りリリースされたことです。同社の方向転換が順調に進んでいることが確認されました」



「インテルは、2024年までに AI があらゆる場所に普及すると認識しており、同社のプロセッサが、インターネットと Windows 等のコンピューターの OS に膨大に存在するすべてのソフトウェア・テクノロジーの基盤となることを望んでいます。これにより、お気に入りの曲を数回クリックするだけで編集し直したり、旅行の写真をコンピューターから簡単かつ迅速に編集したりできるようになります。インテル® Core™ Ultra プロセッサは、あらゆる人をアーティスト、作家、ミュージシャンに変えるでしょう」



「インテルは、見栄えの良い数値データや主張を投げかけるだけでなく、その最新シリコンが可能にする AI ワークロードの種類について、具体的な実例をいくつも提示しています。例えば、レストランでは、利用者の予算や栄養ニーズに応じてメニューを選択できるようになり、製造業では、工場の現場で品質や安全性の問題を検出する新たなシステムを構築できるようになります。インテルのシリコンを搭載した高度な AI は、人間の医師が見逃すような問題を発見できる、より効果的な超音波検査システムの開発にもつながるでしょう」

AI が変革する世界のビジネス

AI がビジネスにもたらすメリット
ビジネスで AI を活用すると、利益を増加させて効率を高めることができます。



データ暗号化



パーソナライズされた学習



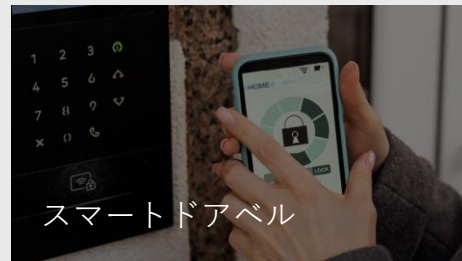
在庫管理



顔認識



ビデオ会議



スマートドアベル



ロボットビジョン



自律走行車



デジタル・アシスタント



AIベースのレンダリング

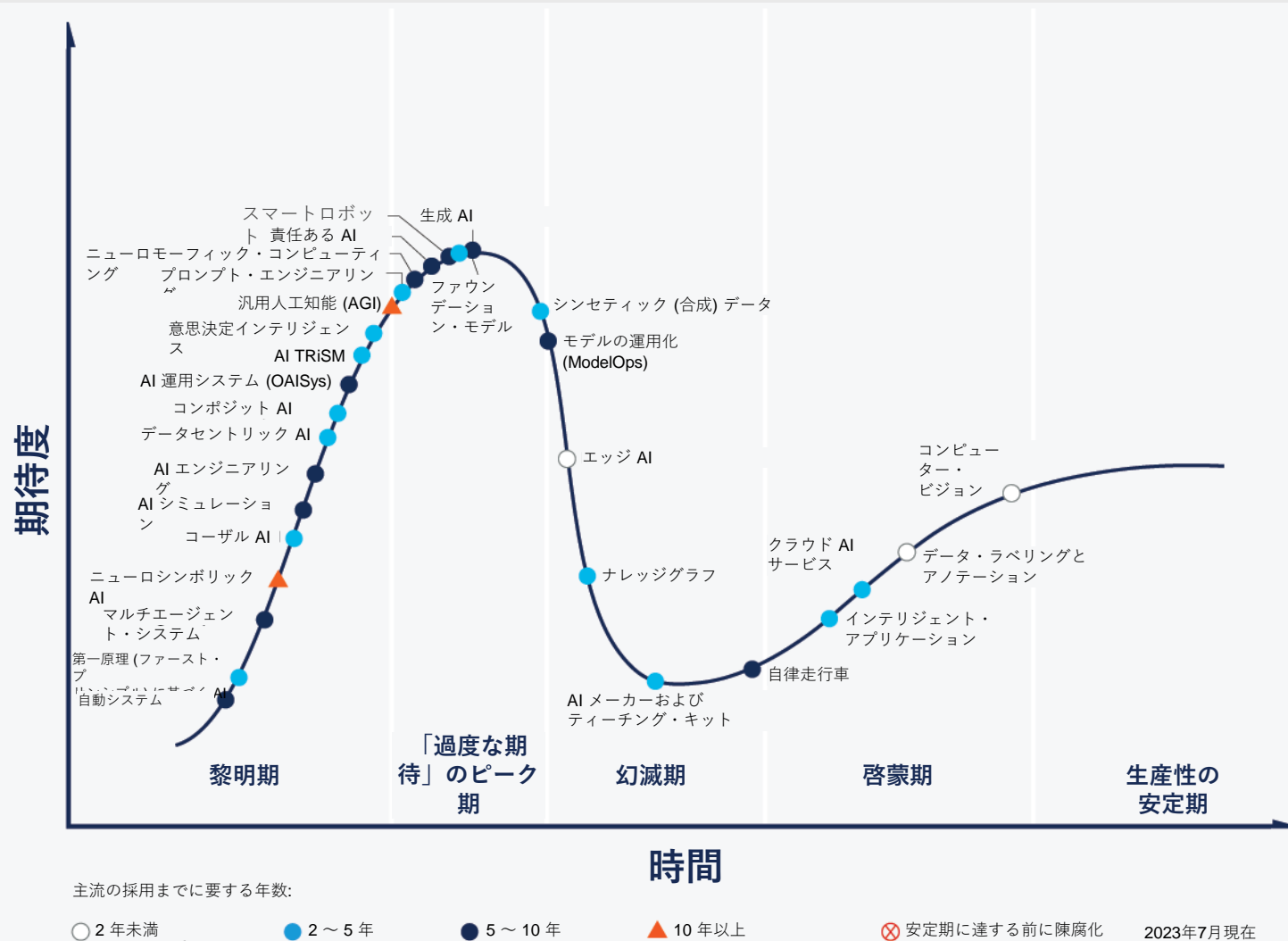


コード生成



音声アシスタント

Gartner - 人工知能 (AI) のハイプ・サイクル



Gartner の「人工知能 (AI) のハイプ・サイクル™」:

2023年」では、不完全なシステムから生じる限界とリスクについて言及する一方で、明らかに革新的なメリットをもたらすイノベーションと手法について特定しています。

また、AI 戦略においては、最も信頼性の高い投資先を検討することが重要です。



「こうしたイノベーションを早期に導入することで、ビジネス・プロセス内で AI モデルを利用する際に生じる問題を緩和し、競争優位性を大きく向上させることができます」

Gartner ディレクター・アナリスト、Afriz Jaffri

AI は急速に進化している

基盤となるデータ・テクノロジー:



3,000 億
米ドル

2026年までに世界の生成 AI 支出額は 3,000 億米ドルを超える見込み

あらゆる場所に
AI

2026年までに

超

50%

の企業で管理されるデータは、データセンターやクラウドの外側で作成・処理されるようになる

50%

のエッジ導入には AI が関係する

58%

の大手上市企業の CEO は、AI に積極的に投資している

インターネットのように
破壊的な AI

生成 AI は、2040年までに世界経済に最大 4.4 兆米ドルの価値をもたらすと予測される²

AI 推論により、ムーアの法則を上回るペースでコンピューティング・コストが上昇する

大規模モデルのサイズ (1 兆以上のパラメーターのモデル) の拡張

小型で軽快なモデル (100 億未満のパラメーター) の拡張

¹ <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#key-insights>

² Worldwide Artificial Intelligence Spending Guide (IDC)

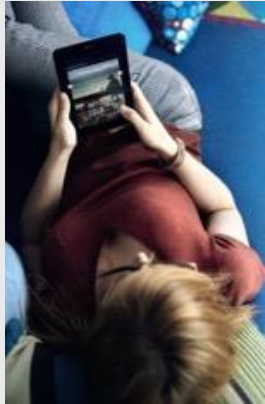
<https://chiefexecutive.net/the-rise-of-the-ai-ceo/>

<https://blogs.gartner.com/andrew-white/2021/07/24/by-2024-60-of-the-data-used-for-the-development-of-ai-and-analytics-projects-will-be-synthetically-generated/>

Gartner©, Hyperscalers Stretching to the Digital Edge (2023年7月)。Gartner は、米国内およびその他の国における Gartner, Inc. および / またはその関連会社の登録商標およびサービスマークであり、許可を得て使用しています。All right reserved.

2022年には 5% だったエッジ・コンピューティングの導入が、2026年までには少なくとも 50% が機械学習 (ML) を含むようになる (「Building Edge Computing Strategy」2023年4月)

ビジネスチャンス ユースケースの例



教育

教師の
アシスタント

学生の学習支援

保護者用
チャットポータル



医療

創薬

ドクター・
コパイロット

患者家族用
チャットボット



財務

アルゴリズム取引

顧客
ポートフォリオ・
アシスタント

リスク/信用評価



小売

製品
プロモーション

顧客
インターフェイス
と感情分析ツール

画像ショッピング
支援



政府機関

行政サービスの
チャットボット

文書検索の要約

リアルタイム
言語翻訳



エネルギー

消費予測

運用
パフォーマンス

エネルギー取引
アシスタント



自動車

自動車開発

多言語の
自動車支援

サプライチェーン
の最適化

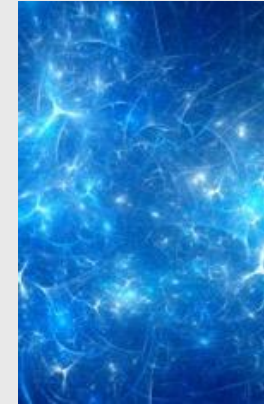


製造

ファクトリー・
オートメーション

予測メンテナンス

精密農業



電話会社

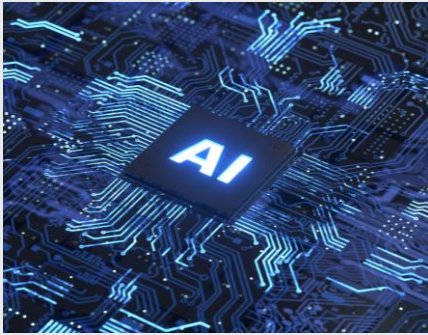
パーソナライズ
された
カスタマー
サービス

ネットワーク・
オートメーション

運用
パフォーマンス

今日の AI には、どのような課題があるでしょう

インテルのパーカプーになる理由



GPU の提供状況

世界的な GPU 不足に
インテル CPU で代替

[Naver の AI サーバー変更](#)は、世界の IT 企業が NVIDIA の GPU 値上げや、世界的な GPU 不足に不満を募らせている中で実現しました。



ベンダー ロックイン

オープンソースの標準
ベースのソフトウェアによる

ベンダーロックインの回避

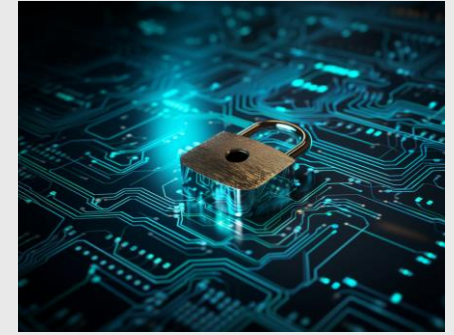
[インテルは、業界標準のすべてのオープン・フレームワーク](#)やライブラリーを取り入れて、最高のパフォーマンスを發揮できるよう最適化し、インテルのテクノロジーですぐ利用可能な高品質のエクスペリエンスを確実に届けます。



コスト

インテルが第 4 世代インテル® Xeon® プロセッサーで、より優れた価格とパフォーマンスを
実現

[実際の業務アプリケーションにおいて](#)、インテルは、パフォーマンスの向上、低価格、AI 推論向けのバランスに優れたプラットフォームを提供することで、業界に破壊的革新をもたらし、AI を民主化しています。



安全な AI

インテルが最も包括的な
セキュリティー・
ポートフォリオを提供

[インテルのセキュリティー機能](#)により、ワークロードに適した信頼境界を設定し、機密データ、コンテンツ、ソフトウェア IP を高度な攻撃、改ざん、盗難から保護します。

インテルが AI をあらゆる場所に導入する方法

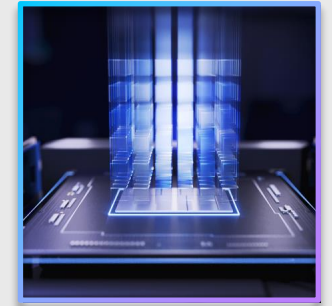
地球上のあらゆる人の生活を向上させ、世界を変えるテクノロジー

インテルの独自の価値

- オープンなアプローチ
- ハードウェアとソフトウェアに関する専門知識
- エコシステム
- 実行



インテルは、AI を実現するテクノロジーの幅広いポートフォリオ、AI を強化する将来のイノベーションに関する独自のビジョン、オープンなエコシステムに対する比類のないサポートにより、**AI をあらゆる場所に導入**し、それがすべての人に利益をもたらすよう支援しています。



- ワークロードのあらゆる領域に AI を拡張し、個人や組織がアクセスできるようにする
- データセンター、クラウド、PC、エッジにまたがる多種多様な AI ワークロードを顧客が自信を持って保護できる、異種混合のアーキテクチャー、オープンな標準、ソリューション

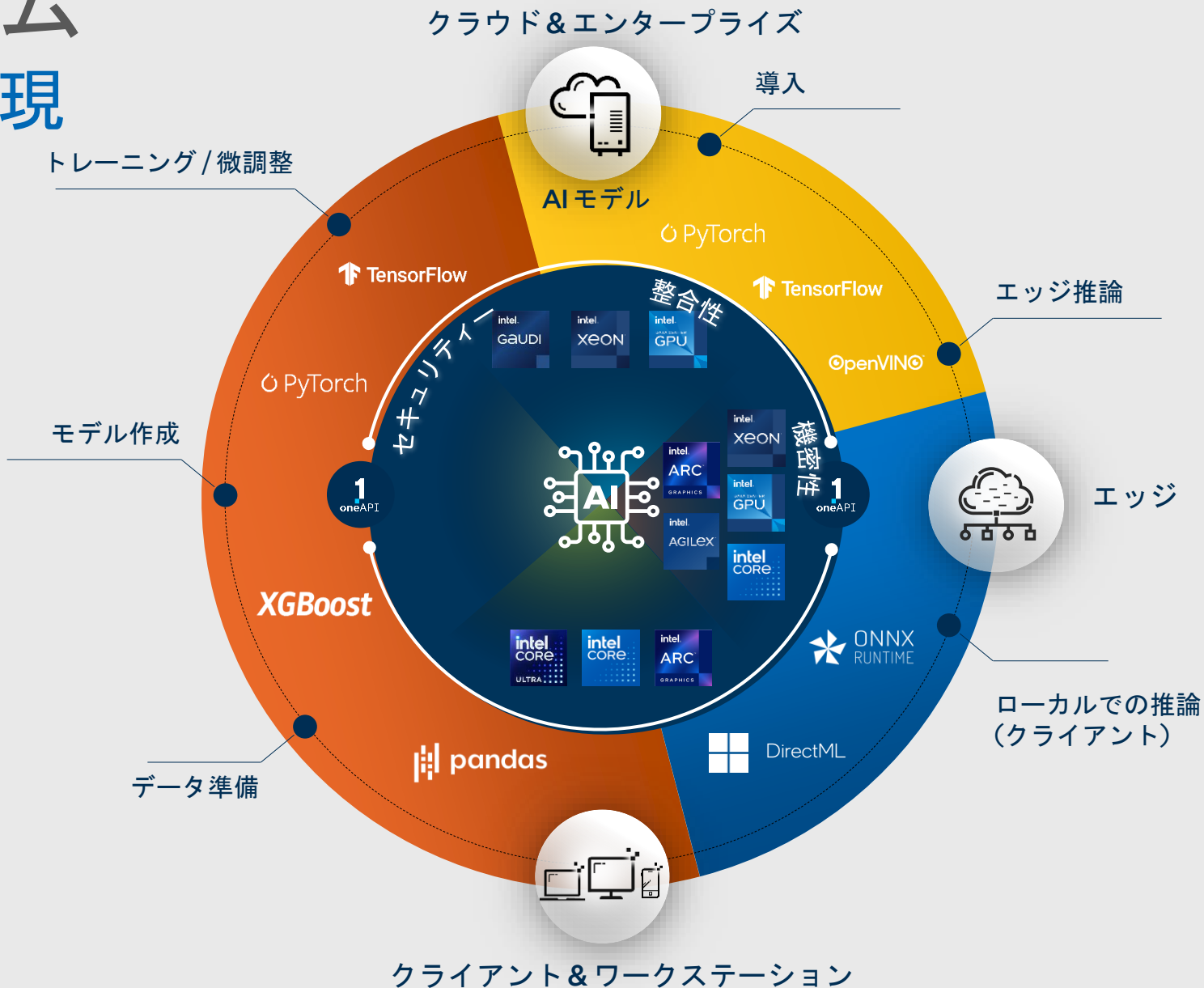


AI コンティニューアム

AI をあらゆる場所で実現

インテルは、
AI をあらゆる場所に導入
 し、AI 導入への道のりの
 あらゆる段階でビジネス
 をサポートする信頼でき
 るパートナー

データセンターから
 クラウド、ネットワーク、
 クライアント、エッジに
 至るまで



注: インテル® Core™ Ultra プロセッサは Meteor Lake 以降、NPU 低消費電力推論エンジンを統合しています。

インテルによる 責任ある AI



環境のためのソリューションの創出

AI テクノロジーを活用することで、研究者は私たちが生活するこの世界の環境の仕組みについて理解を深め、より良い未来を築くためのソリューションを開発できます。



医療の進歩

現在 AI は、患者ケアの向上から疾患を予防するための研究まで、医療やライフサイエンスの分野で一般的に使用されています。



アクセシビリティの強化

障害を持つ多くの人々にとって、自立生活と自主性は課題となる可能性があります。AI は、日常生活に存在する障壁に対して代替ソリューションを提供する製品を生み出すことで、障害を持つ多くの方の生活を変えることに役立っています。



教育へのアクセスの拡大

インテルは、「AI for Youth」(青少年向け AI プログラム) や「AI for Future Workforce」(AI 人材育成プログラム) などのプログラムを活用して、学生をデジタル革命に備えさせるため、世界的な AI のスキルギャップに対応する取り組みを行っています。



安全性の向上

自動運転車による運転の実現から児童労働搾取の軽減まで、AI テクノロジーはより安全な社会作りに役立っています。

あらゆる段階でインテルが AI を推進する方

インテルは、**AI をあらゆる場所に導入**し、AI 導入への道のりのあらゆる段階
で

皆様をサポートする信頼できるパートナー



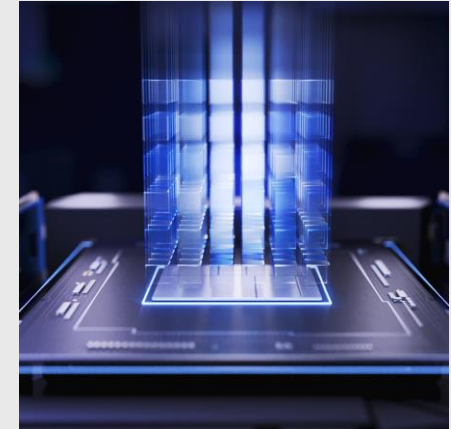
価値を最大化



どこにでも導入



安全な状態の維持



エコシステムへの
投資

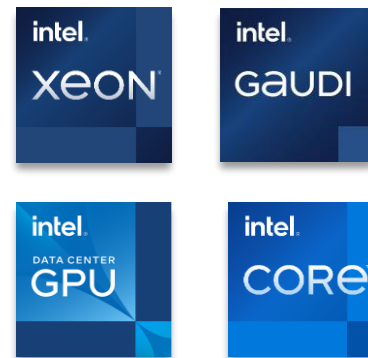


価値を最大化

インテルのオープン AI アプローチが
AI ビジネスのニーズに適している理由

ベンダーロックインを回避
オープンソースの標準ベース
のソフトウェア

インテルのハードウェア・ポートフォ
リオを活用
AI のユースケース向けに最適
化



明日の AI のためにソフトウェアとオープンな標準により最適化されたハードウェアで、
クライアントからエッジ、データセンター、クラウドに至るまで、新たな機会を創出



インテルの AI 戦略

AI イノベーションを加速するためにインテルがもたらすもの

AI アプリケーションとソフトウェア

オープン

能率的

アクセス可能

新しいアルゴリズム

大規模でのパフォーマンス向上



インテル® Tiber™ AI
クラウド

ハイブリッド AI

OpenVINO™



オープンな標準とソフトウェアの使用

データセンター

スケーラブルなシステム

アクセラレーター、Xeon®

ネットワーキング

オープンス・タンダード

ネットワーク・
インフラストラクチャー

クライアントとエッジ

AI PC

NPU / GPU / CPU

あらゆるプラットフォームに内蔵・高速化

高度なハイパフォーマンス・テクノロジー

オープン AI システム・ファウンドリー

高度で責任あるプロセスを使用

倫理的リーダーシップの基盤

AI データの信頼性と安全性を維持



価値を
最大化

インテル® AI ポートフォリオ

AI コンピューティングのあらゆるニーズに最適化された
ハードウェアとソフトウェアをご活用ください

オープンなソフトウェア環境

1 **oneAPI** **OpenVINO™**

PyTorch XGBoost MODIN

TensorFlow ONNX RUNTIME deepspeed

ディープラーニング
アクセラレーション



専用のディープラーニング・トレーニングと推論

汎用の
アクセラレーション

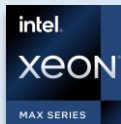


クラウドゲーム、VDI、メディア分析、
リアルタイム高密度映像

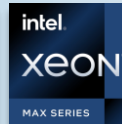


並列コンピューティング、HPC、HPC 向け AI

汎用



リアルタイム、中程度のスループット、
低レイテンシー、スパース推論



中規模から小規模のトレーニング
と微調整



エッジとネットワークの AI 推論



クライアントでの推論



インテル® AI ソフトウェア・ポートフォリオ

エンジニアのデータ

モデルの作成

最適化と導入

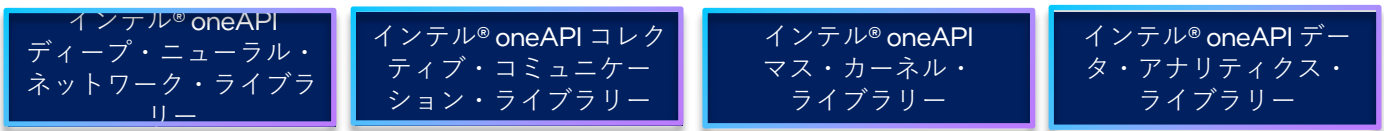


大規模なデータ分析†



マシンラーニング & ディープラーニングのフレームワーク、最適化と導入ツール†

1 oneAPI



CPU、GPU、その他のアクセラレーター向けのオープンなクロスアーキテクチャー・プログラミング・モデル

クラウド & エンタープライズ



クライアント & ワークステーション



エッジ



エンドツーエンドのデータサイエンスと AI を加速



インテル® Tiber™ AI クラウドとインテル® Developer Catalog
最新のインテルのツールとハードウェアを試して、最適化された AI モデルにアクセス

cnvrg.io

フルスタックの ML オペレーティング・システム

インテル®

Geti™
アノテーション・トレーニング / 最適化プラットフォーム

Hugging Face

インテルの最適化と微調整レシピ、最適化された推論モデル、モデルの提供

注: スタックの各レイヤーのコンポーネントは、予想される AI の使用モデルに基づいて、その他のレイヤーの対象となるコンポーネントに最適化されています。右側の列にあるソリューションでは、すべてのコンポーネントが活用されるわけではありません。

†このリストには、インテルのハードウェアに最適化された、一般的なオープンソースのフレームワークが含まれています。



リファレンス・キットで AI 開発を加速

最適化された AI リファレンス・キットが、開発者とデータ・サイエンティストの
イノベーション迅速化を支援

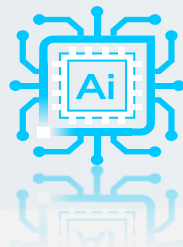
オープンで規格に準拠した、ヘテロジニアス・プログラミング・モデルである [oneAPI](#) と、[インテル® AI アナリティクス・ツールキット](#) や [インテル® ディストリビューションの OpenVINO™ ツールキット](#) などの、インテルのエンドツーエンドの AI ソフトウェア・ポートフォリオに基づいて構築されたリファレンス・キットにより、AI 開発者はアプリケーションへの AI 導入プロセスを効率化でき、既存のインテリジェント・ソリューションを強化して導入を加速します。

結果的に、従来のモデル開発ワークフローと比較して、より短く生産的なワークフローにより、パフォーマンスが改善されることが実証されています

企業の会話型 AI チャットボットとのインタラクションを設定するように設計された **AI リファレンス・キット** を使用すると、ユーザーは oneAPI の最適化により、バッチモードで **推論を最大 45% 高速化** できます



ライフサイエンス向け視覚的品質管理検査を自動化するために設計された **AI リファレンス・キット** は、oneAPI の最適化により、視覚的欠陥検出において **トレーニングを最大 20% 高速化し、推論を 55% 高速化** します



開発者が公共設備のアセットの健全性を予測し、より高いサービスの信頼性を実現するために、**AI リファレンス・キット** が用意されており、予測精度を **最大 25% 向上** させています





安全な状態の
維持

安全な状態の維持

内蔵のセキュリティー機能により、AI イニシアチブを保護し、規制に準拠

セキュリティー
機密データとモデルを保護



コンプライアンス
セキュリティーと
プライバシーの規制
に準拠




機密性
プライベートなデー
タを公開することな
く、
複数パーティーの AI





インテルが最も包括的な セキュリティー・ポートフォリオを提供

インテル® ソフトウェア・
ガード・エクステンション
ズ
(インテル® SGX)



アプリケーションの分離




インテル® トラスト・
ドメイン・エクステンシ
ョンズ (インテル® TDX)



仮想マシンの隔離



インテル® Trust
Authority



マルチクラウドとハイブリッ
ド・
クラウド向けの独立した
信頼性検証サービス

ソフトウェア・ソリューション、クラウド、OEM、システム・インテグレーターのエコシステ
ム

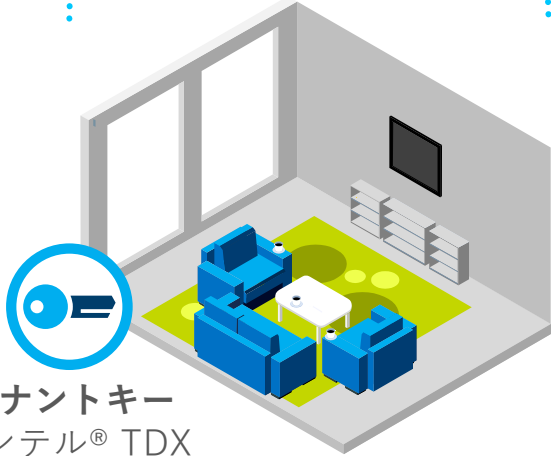
インテルのセキュリティー・ファーストの開発とライフサイクル・サポート



メモリー内のデータの保護

今日: メモリーを盗み見ることができれば、データの復号化に使用される秘密鍵など、そこを通過するすべての情報を見ることができます。

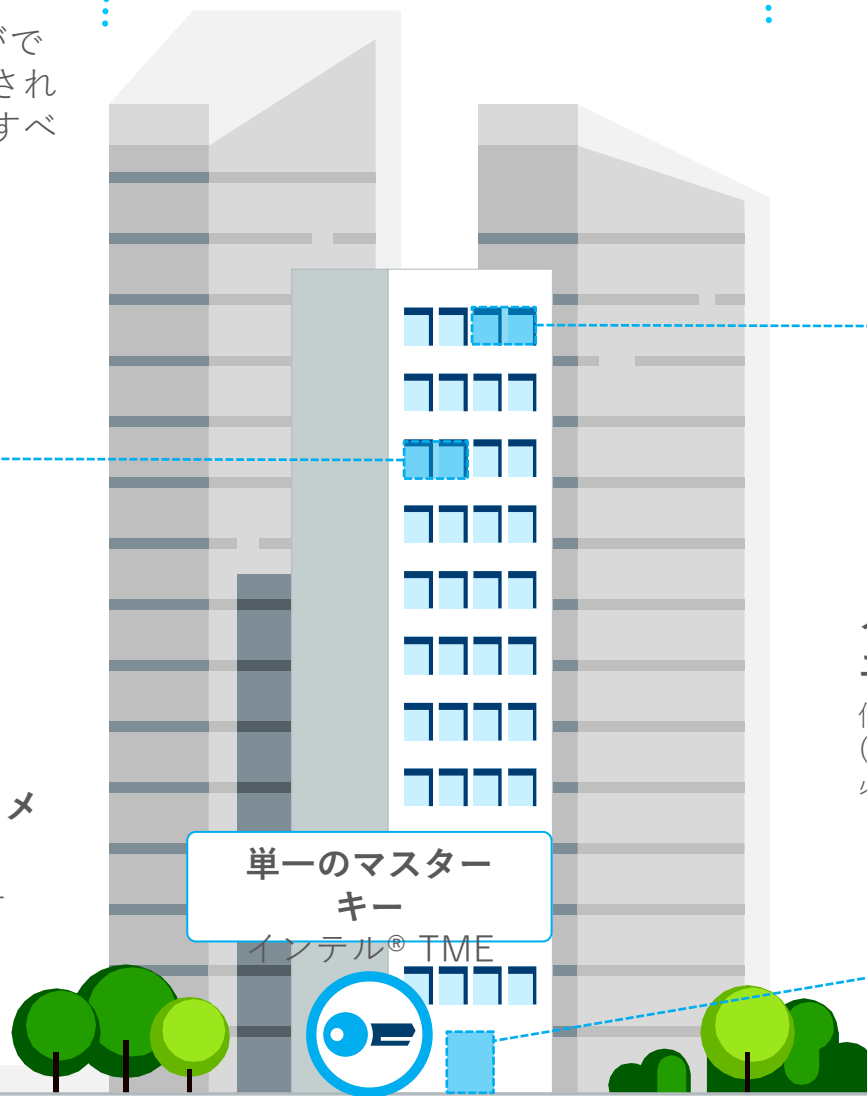
例えるなら「VM」



インテル® TDX = インテル® トラスト・ドメイン・エクステンションズ

各 VM スペースを個別に暗号化するための個別のキー (機能の認識には OS/VMM のみが必要)

例えるなら「システムメモリー」



例えるなら「アプリケーション」



インテル® SGX = インテル® ソフトウェア・ガード・エクステンションズ

個々のアプリケーション・データ・スペースの分離 (アプリケーション・コードの変更または抽象化インターフェイスが必要)

インテル® TME = インテル® トータル・メモリー・エンクリプション

単一のキーでシステムメモリー全体を暗号化 (OS/App の変更は必要ありません)



インテル® セキュリティー・エンジンで イノベーションを加速し、データ保護を強化

インテル® Xeon® スケーラブル・プラットフォームによるコンフィデンシャル・コンピューティング

データの機密を保ちながら、データを活用

インテル® Xeon® CPU に搭載されたインテル® セキュリティー・エンジンは、データの機密性とコードの整合性を確保しながら、パフォーマンス

を維持:
 [製品概要](#)

インテルの採用により、重要なビジネスの成果につながる、より良いインサイトを

得ることができる:
 [ビジネス概要](#)

コンフィデンシャル・コンピューティングの導入に役立つ

[インテル® SGX](#) および [インテル® TDX](#)



AI エコシステムの実現

AI 開発者が好む最新のソフトウェア・ツールを使用し、最適化されたパフォーマンスで、
新しい機会と重要なビジネスの成果を促進

オープン
プログラムが可能

選択肢
互換性

信頼性
AI 推論

安全な AI
安全なワークロード
安全なモデル
安全な使用中データ

スケール
開発とテスト

インテル® Tiber™ AIクラウド

cloud.intel.com

小規模、中規模、大規模モデル

フルシステム、フルクラスター

最新のインテル CPU、アクセラレーター
およびソフトウェア

AI 向けのオープンで高速化されたコネクテッド・コンピューティング

マルチベンダー マルチアーキテクチャー

ハードウェア/アーキテクチャー

AIをあらゆる
場所に導入

インテル® Xeon® プロセッサーで AI を実行

クラウド &
エンタープライズ



エッジ





第5世代インテル® Xeon® プロセッサー： AI 向けに設計されたプロセッサー

すべてのコアに AI アクセラレーションを搭載した第5世代インテル® Xeon® プロセッサーにより、お客様がディスクリット・アクセラレーターを追加する前に、要求の高いエンドツーエンドの AI ワークロードに対応

AI 推論の
パフォーマンス向上

最大 **42%**
前世代との比較¹

一般的な
コンピューティング・
パフォーマンスの向上

平均 **21%**
前世代との比較¹

自然言語処理の高速化

最大 **23%**
前世代との比較¹

インテル コーポレーション 主
席副社長 データセンター & AI
事業本部 本部長
サンドラ・リベラ

「AI 向けに設計された第5世代インテル® Xeon® プロセッサー・ファミリーは、クラウド、ネットワーク、エッジの各ユースケースで AI 機能を導入するお客様に、より高いパフォーマンスを提供します。顧客、パートナー、開発者エコシステムとの長年の協力により、第5世代インテル® Xeon® プロセッサーは、TCO を削減した迅速な導入と拡張を可能にする、実証済みの基盤の上に、リリースされています」

詳細情報

[ウェブサイト](#)

[製品概要](#)



インテル® Xeon® プロセッサー: 実世界の AI アプリケーションにおける CPU パフォーマンスのリーダーシップ

実際の業務アプリケーションにおいて、インテルは、以下の方法で、優れたパフォーマンス、低価格、バランスの取れた

プラットフォームを提供することで、業界に破格的変革をもたらす AI の民主化を実現します。



- データ・ローカルティに役立つキャッシュ容量の増加と、より大きな問題を解決できる大容量のメモリー



- コア周波数の向上、複数のスカラーポート、アウトオブオーダー実行により、シングルスレッドまたはマルチスレッドでありながらスカラーである計算処理の高速化を支援



- 非 DL ベクトル演算を支援するインテル® アドバンスト・ベクトル・エクステンション 512 (インテル® AVX-512)



- AI アクセラレーションを内蔵ハードウェアでサポートするインテル® アドバンスト・マトリクス・エクステンション (インテル® AMX)

[技術記事の全文](#)

[インフォグラフィック](#)



[GPU 神話の虚像: 内蔵アクセラレーター搭載 CPU が AI に革命を起こす方法](#)

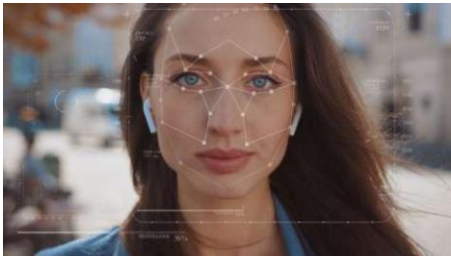


AI 推論用アクセラレーターを搭載した第 4 世代 インテル® Xeon® スケーラブル・プロセッサー・ ファミリー

インテル® AVX-512、インテル® AMX などのアクセラレーターは、パフォーマンスの向上、レイテンシーの削減、

インテル® アドバンスド・マトリクス・エクステンション
(インテル® AMX)

ディープラーニングのトレーニングと推論を大幅に高速化。
自然言語処理、推奨システム、画像認識などの
ワークロードに最適



[ウェブサイト](#) | [ソリューション概要](#)
[ビデオ](#) | [ユーザーガイド](#)

インテル® アドバンスド・ベクトル・エクステンション 512
(インテル® AVX-512)

古典的機械学習や、データの準備などの
エンドツーエンドの AI ワークフローにおける
その他のワークロードを高速化



[ウェブサイト](#) | [ソリューション概要](#)
[ビデオ](#) | [ユーザーガイドとダウンロード](#)



インテル® AMX 搭載第 4 世代インテル® Xeon® スケーラブル・プロセッサー・ファミリーが、AMD EPYC を上回る性能を発揮

より速く、よりパーソナライズされた AI で収益向上を促進し、顧客体験の向上を実現



ビジネス上の意思決定をよりの確にして、収益の拡大を促進する



顧客の維持と獲得を改善する



エンゲージメントとコンバージョン率を向上させる



業務で反復的な作業、コスト、時間を削減する



VS



インテル® アドバンスド・マトリクス・エクステンション (インテル® AMX) 搭載第 4 世代インテル® Xeon® スケーラブル・プロセッサー・ファミリーが、
どのように AMD EPYC より優れているかをご確認ください



AI ワークロード: インテル® Xeon® スケーラブル プロセッサ・ファミリーで VMware を実行する

vmware®



インテル® アドバンスト・
マトリクス・エクステンション
(インテル® AMX)

「内蔵 AI アクセラレーションを搭載した CPU を使用して、データ準備、トレーニング、最適化、推論といった、エンドツーエンドの AI パイプライン全体を実行できま

「AI/ML ワークロードのパフォーマンスを向上させるには、CPU の AMX 命令に AI/ML の動作の一部を行わせて、高価で調達が困難な GPU の必要性を下げる方法があります」



VMware スタッフエンジニア、Earl-Ruby 氏による [記事全文](#)

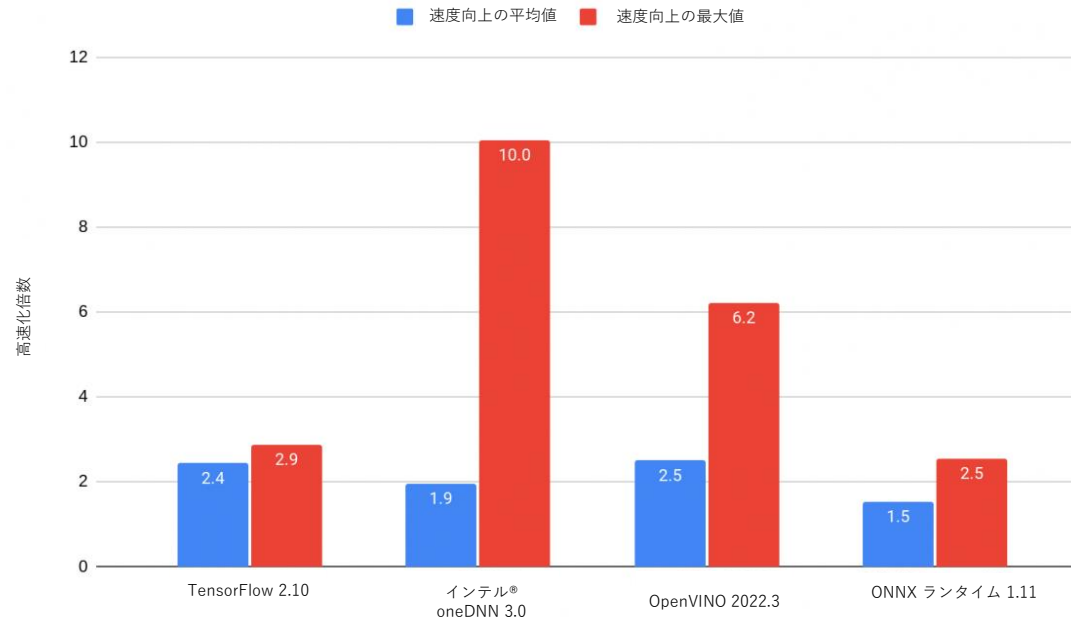


AI ワークロード: Red Hat の実行 インテル® Xeon® スケーラブル・プロセッサ・ ファミリー

Red Hat Enterprise Linux のパフォーマンスが大幅に向上
第 4 世代インテル® Xeon® スケーラブル・プロセッサ・ファミリー

AMX

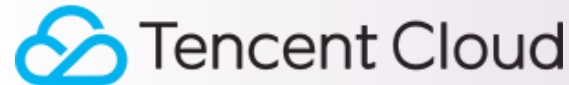
4P Cooper Lake と比較した、2P Sapphire Rapids の Phoronix テストスイートの速度向上の倍率



測定結果では、第 4 世代の高速化の割合は平均 1.5 倍~最大 10 倍¹



導入事例



クラウド・コンピューティング・サービスの
検索エンジン

課題

大規模なクエリーを処理し、検索結果で迅速に回答する方法

ソリューション/結果

Tencent は最適化された BERT モデルを使用して、**より良いサービス体験を提供して TCO 削減**の支援を行うことができる

インテル製品

第 4 世代インテル®
Xeon® + インテル® AMX

詳細

[導入事例](#)



大手リテール・テクノロジー企業

費用対効果の高いビジョン AI サービス

Meituan は、オンラインリソースの全体的な効率を **3 倍以上**向上させ、**サービスコストを 70% 削減**

第 4 世代インテル®
Xeon® + インテル®
AMX + インテル® IPP
+ PyTorch 向けインテ
ル® エクステンション
(インテル® IPEX)

[導入事例](#)



医用画像処理

放射線治療専門家の
作業効率の向上

放射線治療の専門家を AI ベースの自動輪郭作成技術を用いてサポートすることで、ワークロードの効率を向上し、一貫性を高め、スタッフの負担を軽減し、付加価値の高い作業に集中できるよう支援

第 4 世代インテル®
Xeon® + インテル®
AMX + OpenVINO™

[導入事例](#)
[ビデオ](#)



大手クラウド・コンピューティング・プロバイダー

アドレス浄化サービスのパフォーマンスを向上

エンドツーエンドのパフォーマンスの高速化は、**物流、eコマース、エネルギー、小売、金融**の Alibaba 顧客にとって、**ビジネス成果に直結し、内蔵アクセラレーターの使用により、Alibaba が TCO の抑制をサポート**

第 4 世代インテル®
Xeon® + インテル®
AMX + インテル®
oneDNN

[導入事例](#)



インテルの AI テクノロジーに関するお客様の声



「数週間分のセットアップ
時間を短縮しました」

「私たちにとって、イン
テル® Xeon® プロセッ
サーは、テクノロジーの
導入方法の土台です。私
たちはインテル® Xeon®
CPU でのみ稼働しており、
VM や専用のオンプレミ
ス・ベアメタル、クラウ
ドなど、あらゆる場所
での稼働する能力を有して
います。」



[導入事例](#)

SIEMENS

35 倍 自動輪郭作成
アルゴリズムの AI 推論時間が高
速化 (前世代との 比較)¹

20% 消費電力を削減
(前世代との 比較)²



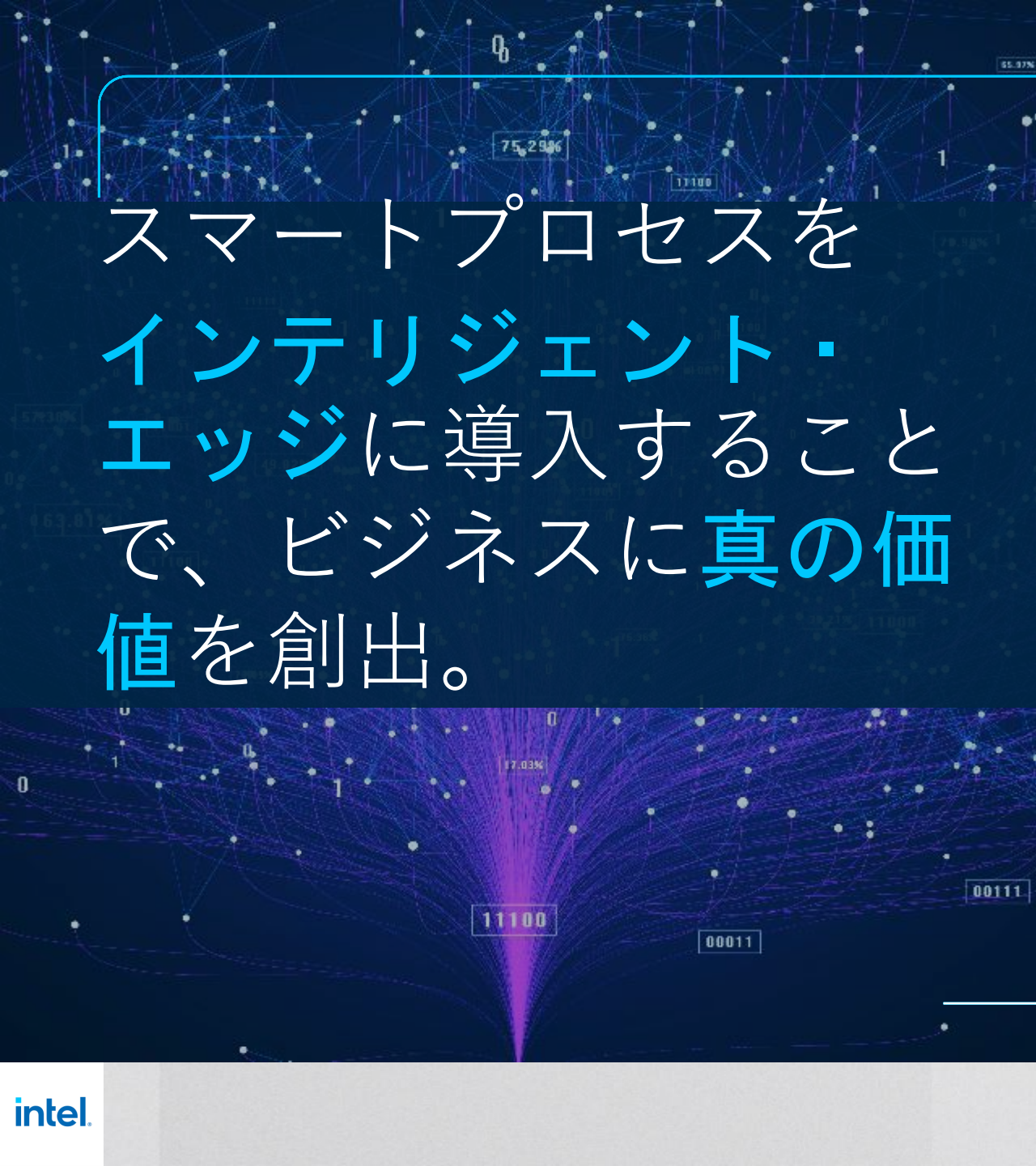
[導入事例
ビデオ](#)

エッジ AI

エッジ



AIをあらゆる
場所に導入



スマートプロセスを
インテリジェント・
エッジに導入すること
で、ビジネスに**真の価値**
を創出。

エッジ・コンピューティングと AI を
連携させた力を活用することで、
ビジネスの成果を高め、
顧客エクスペリエンスを向上。



エッジでのデータの処理

2025年までに、企業が生成するデータの75%がエッジで作成、処理されるようになる¹

低レイテンシー

エッジ・コンピューティングでは、保存と処理前のクラウドとのデータのやり取りが不要なため、インサイトまでの時間が短縮され、効率が向上します。

リスクの低減

エッジのデータはIoTデバイス自体に格納、処理されるため、イベントに対するリアルタイムで迅速な対応が可能になり、ビジネスリスクの軽減とセキュリティの強化が実現します。

コストの削減

エッジ・コンピューティングは、データをエッジに保持し、よりコスト効率の高いストレージと処理を実現すると同時に、ビジネスプロセスを合理化する、より迅速なインサイトを実現しま

エッジでのコンピューティングと処理により、作成された場所^{す。}でデータを活用する機会を創出

¹出典: <https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders>

エッジ



エッジ AI が 業界を変革する ユースケースを 実現

導入事例

- [VSBLTY](#)
- [Convergent](#)
- [Taco Bell](#)



教育

- 教室環境に関する重要なインサイトで、学習環境を強化
- 行動に関するインサイトで、より魅力的なレッスンプランを作成
- AI ベースのビデオ監視で学校のセキュリティを向上



エネルギー

- AI を活用した設備監視で環境への影響を軽減
- 自動監視で電力コストを削減



政府機関

- 人事管理の最適化
- AI ベースのビデオ分析で建物の安全性とセキュリティを強化
- エネルギーの無駄を削減



医療 & ライフサイエンス

- 診断と医療検査でインサイトを得るまでの時間を短縮
- 診断の精度を向上
- 患者ケアの向上



製造

- 品質管理プロセスの向上
- 従業員の健康と安全を確保
- メンテナンス・コストを削減
- 予測メンテナンスの有効化



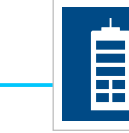
小売

- 魅力的なパーソナライズされた広告で店舗へのトラフィックを増加
- インテリジェントなプロモーションで売上を向上



輸送

- ロジスティクスの精度と効率を向上
- インテリジェントなパッケージ取り扱いで送料と返品を削減



接客業

- レストランの顧客エクスペリエンスのパーソナライズと改善
- QSR の注文と待ち行列の合理化
- 調理を最適化して廃棄物を削減

▶ インテルはエッジ AI を活用して多様な環境での体験を向上させています



内蔵型 AI とセキュリティー機能で クリティカルなエッジ・ワークロードを高 速化

第 4 世代インテル® Xeon® スケーラブル・プロセッサー・ファミリー
第 3 世代とのパフォーマンスの比較

1.33 倍

より高い性能

3.01 倍

画像分類にインテル® AMX
を使った場合に AI 推論
パフォーマンスが向上

4.25 倍

物体検出にインテル® AMX
を使用した場合に AI 推論
パフォーマンスが向上

IoT エッジ向け第 4 世代インテル® Xeon® スケーラブル・プロセッサー・ファミリーは、驚異的なパフォーマンス、メモリー、I/O、ワークロードの統合をサポートするリソース管理機能、エッジでのディープラーニングのトレーニングと推論を可能にする新たな AI 命令セットを提供

[詳細情報](#)

[AI in Production の成功事例](#)

AIをあらゆる
場所に導入

PCへのAI導入

クライアント&
ワークステーション





ユースケース: PC への AI 導入

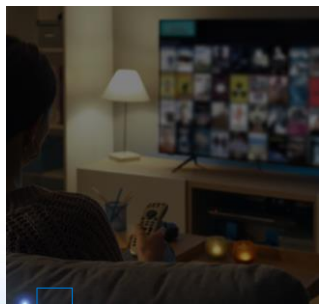
クリエイター: 写真と動画の検索と編集

自動化され、迅速化された検索で、より高速で自然なフィルター、より高品質なプレビュー、エクスポート時間の短縮を実現。



コラボレーション/ ストリーミング

次世代のビデオ会議、ストリーミング、コラボレーション、バッテリー持続時間の維持を実現する新しい AI 機能。



メインストリームのゲーム

ゲーム内の臨場感を高める 3D アニメーション、文字起こし、チャット翻訳などの新しい AI 機能。



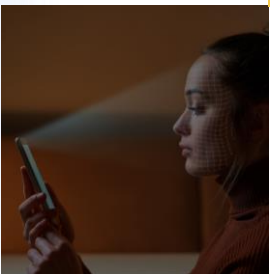
PC への AI 導入

生産性

執筆、創作、コーディング、テキストと文法予測などのオフライン機能の AI アシスタント。

アクセシビリティ

多様なユーザーのニーズに対応する AI 支援のオーディオビジュアル機能により、PC での制作が容易になり生産性が向上。



クリエイター: テキストから画像への変換

マーケティング、広告、デザインなど、わずかな言葉で画像を生成できる新しい AI エフェクトと機能。

「AI が可能性を解き放つ」



AI PC 時代の到来を告げる インテル® Core™ Ultra プロセッサー

AI PC の心臓部となる最新プロセッサーは、オペレーティング・システムとアプリケーションの全体で、AI 機能を活用

Intel 4 プロセス・
テクノロジーで構築され
た
最初のプロセッサー
40 年間で最大の
アーキテクチャー・シフ
ト

最大 8 つの Xe コアを
搭載する内蔵
インテル® Arc™ GPU3
前世代と比較して最大
2 倍のグラフィックス・
パフォーマンス¹

インテル® AI Boost -
低消費電力で長時間動作
する AI ワークロード向けに
開発されたインテル
最新 NPU
前世代と比較して最大 2.5
倍電力効率が向上¹

インテル コーポレーション 主席副社長
クライアント・コンピューティン
グ事業本部 本部長 ミッシェル・ジョ
ンストン・ホルトハウス

「インテル® Core™ Ultra プロセッサーのリリースは、インテルが PC で AI を実現している比類のない規模とスピードを象徴しています。2028年までに、AI PC は PC 市場の 80% を占めるようになると予想され²、インテルはハードウェアおよびソフトウェアのパートナーの広大なエコシステムと連携して、次世代コンピューティングを実現する最適な立場にあります」

詳細情報

[ウェブサイト](#)

[製品概要](#)

¹ ワークロードと構成については、[intel.com/processorcailms](https://www.intel.com/processorcailms) でインテル® Core™ Ultra 7 プロセッサー 165H のパフォーマンスをご覧ください。結果は状況によって変わります。

² 出典: Boston Consulting Group



導入事例: インテル® Core™ Ultra プロセッサに搭載された AI による患者ケアの進歩

CPU 処理の超音波画像処理アプリケーションを通じて、より身近で費用対効果の高い画像処理テクノロジーを実現

状況

Samsung Medison は、医療イノベーションのパイオニアです。同社の超音波画像処理アプリケーションは、最も効果的な患者ケアを実現するために AI を使用しています。

課題

以前、同社のアプリケーションは、競合他社のディスクリット GPU で高速化された前世代のインテル® Core™ プロセッサで実行されていました。

ソリューション

Samsung が、GPU エンジンを搭載した新しいインテル® Core™ Ultra プロセッサをテストしたところ、前世代の CPU + dGPU の組み合わせと比較して、AI のパフォーマンスが大幅に向上しました。インテル® Core™ Ultra プロセッサの搭載により、Samsung Medison は、CPU のみをベースに、次世代の超音波デバイスに高度な AI 機能を提供することができます。

詳細を
確認:
[詳しくは
こちら](#)



intel
CORE
ULTRA



AI PC アクセラレーション・プログラム

AI PC アクセラレーション・プログラムは、独立系ハードウェア・ベンダー (IHV) および独立系ソフトウェア・ベンダー (ISV) と、人工知能 (AI) ツールチェーン、トレーニング、共同エンジニアリング、ソフトウェアの最適化、ハードウェア、設計リソース、技術的専門知識、共同マーケティング、

販売機会などのインテルのリソースを結びつけることを目的としています。

インテルは AI で業界をリード

AI PC 最適化のために
AI ISV パートナー

100+

と連携

2024年を通じて AI により
加速された ISV 機能

300

超

2025年までに AI アクセラレーター
を搭載するプロセッサ

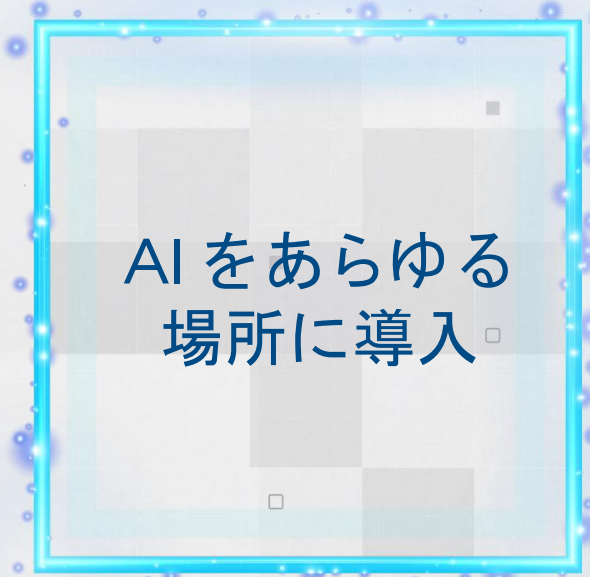
1 億

超

詳細については、今すぐお問い合わせください。

ai.pc.acceleration.program@intel.com

Gaudi2 で AI を実行する



AI をあらゆる
場所に導入



Gaudi2: 基礎モデルの効率的なトレーニングと推論に最適

Gaudi2 は、LLM (GPT) や GAI (Stable Diffusion) といった大規模な基礎モデルの要求を満たすディープラーニングの

要件	Gaudi2
速度	トレーニングと推論の両方で A100 よりも 1.5 ~ 2 倍高速
メモリー	各 Gaudi2 デバイスは、 96GB のオンチップの高帯域幅メモリー を搭載しているため、大規模な基礎モデルをメモリーに簡単に取り込み、トレーニングと導入を実行可能
スケーラビリティ	オンチップに統合された 24x100GbE ポート 、サーバーにある 8 枚のカード間のオールツーオールの直接接続、サーバー内およびサーバー間のオープンな ROCEv2 ベースの通信により、効率の高いスケーリングを実現
使いやすさ	SynapseAI、PyTorch、DeepSpeed を使用し、 最小限のコード変更 でモデルを移行または構築
電力効率	A100 と比較して、 ワット当たりスループットが最大 1.8 倍向上
コスト効率	専用の第 1 世代の Gaudi アーキテクチャーに基づいて、Amazon クラウドで A100 より 最大 40% 優れたコスト・パフォーマンス を実現




Gaudi2: 生成 AI と大規模言語モデルの高速化

Gaudi2 ディープラーニング・アクセラレーターは、ディープラーニングのトレーニングと推論で競争力を発揮し、NVIDIA A100 よりも最大 2.4 倍高速なパフォーマンスを実現¹

Habana Gaudi2 と第 4 世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーが、AI トレーニングで最先端のパフォーマンスと最適なコスト削減を実現²

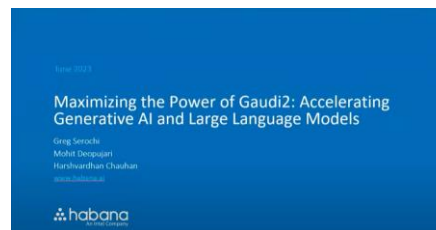
生成 AI と大規模言語モデル (LLM) の可能性を引き出す、Habana® Gaudi®2 AI プロセッサの最先端機能を紹介するインテルのウェビナー録画

GAUDI[®]2
VS
 NVIDIA.



[1 プレスリリース](#)

[2 ニュースルーム](#) [技術記事](#)

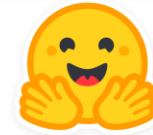


[ウェビナー](#)



ディープラーニング・イノベーション: インテル、Habana Labs、Hugging Face

インテルと Hugging Face の継続的な取り組みの焦点は、最新の
インテル® Xeon® スケーラブル・プロセッサ、Habana Gaudi® および Gaudi®2 プロセッサに
最適化されたトレーニングと推論ソリューションの採用を拡大することです。



Hugging Face

このコラボレーションにより、最先端のディープラーニング・イノベーションがインテル® AI ツールキット
から Hugging Face のオープンソース・エコシステムにもたらされ、将来のインテル® アーキテクチャーの
イノベーション・ドライバーに反映されます。



インテル、Habana Labs、Hugging Face
によるディープラーニング・ソフトウェア
の発展



Habana Gaudi で Transformers
の活用を開始

トレーニングと推論の高速化: Habana Gaudi®2 と NVIDIA A100 80GB の比
較
ベンチマーク



民主化した AI: インテル、Habana Labs、Hugging Face

 habana



Hugging Face

20% 高速 - NVIDIA の A100 と比較した Habana® Gaudi®2 による 1,760 億パラメーターモデルの推論実行速度¹

1.8 倍 の優位性 - Gaudi2 サーバーで一般的なコンピューター・ビジョンのワークロードを実行した場合に、同等の A100 サーバーと比較したワット当たりのスループットの優位性¹

発表を読んで、Fireside Chat をこちらでご覧ください。

[生成 AI に立ちはだかるコンピューティングとサステナビリティの課題に取り組む](#)



ポッドキャスト


[Hugging Face とインテル - 実用的かつ高速、民主化された倫理的な AI ソリューションの実現に向けて](#)



Twitter/X
会話

[民主化された大規模言語モデルが AI 開発を促進する方法](#)

¹ 性能は、使用状況、構成、その他の要因によって異なります。ワークロードと構成の詳細は、「Supermicro L12 Validation Report of Gaudi2 HL-225H SYS-820GH-THR2」(2022年10月20日)を参照してください。



AIをあらゆる
場所に導入

インテル® データセンター GPU マックス・シリーズ
で AI を実行する



インテル® データセンター GPU マックス・シリーズ: 圧倒的なパフォーマンス

インテル最高の性能と密度を誇るディスクリート GPU

インテルの基礎となる GPU コンピューティング・ビルディング・ブロック機能:

- ディスクリート SRAM テクノロジーをベースにした最大 408MB の L2 キャッシュ、64MB の L1 キャッシュ、最大 128GB の広帯域幅メモリー
- 科学的可視化とアニメーションを高速化するために、各インテル® データセンター GPU マックス・シリーズに組み込まれた最大 128 個のレイ・トレーシング専用ユニット
- AI を強化するインテル® Xe^e マトリクス・エクステンション (XMX) に、単一デバイスでベクトルおよびマトリクス機能を実現するディープ・ストリック・アレイを搭載
- oneAPI 標準ベースのマルチアーキテクチャー・プログラミングとツールにより、パフォーマンスと生産性を向上し、独自のプログラミング・モデルのロックインを克服
- 優れたパフォーマンスのハイライト:
 - 最大 12.8 倍のパフォーマンス向上 - インテル® oneAPI ツールにより最適化された 6 つのインテル® データセンター GPU マックス・シリーズにカーネルをオフロードすることで、インテル® マックスシリーズの CPU で実行する LAMMPS (大規模原子 / 分子超並列シミュレーター) ワークロードのパフォーマンスが、第 3 世代インテル® Xe^e プログラミング・モデルと比較して向上

インテル® データセンター GPU マックス・シリーズは、AI および HPC で使用されるデータ集約型コンピューティング・モデルで画期的なパフォーマンスを発揮するように設計されています。インテル® データセンター GPU マックス・シリーズは、SoC の構築において、
より高い柔軟性とモジュール性を実現します。

1

oneAPI

インテル® マックス・シリーズ製品ファミリー全体は、oneAPI で統一され、共通のオープンな標準ベースのプログラミング・モデルによって、生産性とパフォーマンスを高めます。

oneAPI に最適化されたディープラーニング・フレームワークと機械学習ライブラリーを使用して、デベロッパーはデータ分析と機械学習ワークフローのドロップインでの高速化を実現することができます。



導入事例: インテル® データセンター GPU マックス・シリーズ搭載の Aurora スーパーコンピューター

世界で最も困難な問題を解決 - より迅速に



米国エネルギー省所管のアルゴンヌ国立研究所 (ANL) が有する **Aurora** スーパーコンピューターは、1 エクサフロップ以上の持続倍精度性能と 2 エクサフロップ以上のピーク倍精度性能を特徴とする業界初のスーパーコンピューターの 1 つとなることが期待されています。また、Aurora は 1 万以上のブレードと 6 つのインテル® データセンター GPU マックス・シリーズ、2 つの Xeon Max CPU を搭載し、マックス・シリーズの GPU と CPU を単一のシステムに組み合わせた能力を初めて発揮する予定です。

[マシンラーニング用の Aurora ブレードのデモ](#)

実施すること

実施すること

教育



AIのニーズに対してインテルのテクノロジーを利用する方法と、より多くのビジネス獲得に役立つインテル® Xeon® 製品ラインの範囲を理解する

詳細はこちら

[AIトレーニング・アセット](#)

関与



技術分野に関する打ち合わせを始める

テクニカル・ディスクロージャーをご希望の方は、メールで

ご予約ください。

cloud.insider.program@intel.com

インテル® パートナー・アライアンスは どう支援できるのか

インテル® パートナー・アライアンスに登録する

インテル® パートナー・アライアンスのメンバーになると、グローバル市場への参入、高度なトレーニング、プロモーション・サポートなど、御社のニーズに合うメンバー限定のビジネス確立の機会を得られます

トレーニングと コンピテンシー



インテル® パートナー・ユニバーシティに入会すると、知識を身につけるための先進テクノロジーに関する専門トレーニング、コンピテンシー・プログラム、そして特典が提供されます

マーケティング・リ ソース



インテル® ソリューション・マーケットプレイスとインテル® パートナー・マーケティング・スタジオへの参加は、御社製品やサービスに対する

需要を増やす助けになります。
またメンバーでない場合は

今すぐ登録

価値ある報奨



対象となる活動でポイントを獲得し、メンバーのステータスを向上させ、
ビジネス確立のためのさらなるリソースにアクセスしましょう

メンバーシップの特典

ポイントを獲得



インテル® パートナー・アライアンス
の
中で最も人気があり、際だった特典の
1つは、インテルがパートナーに対し
て付与
するポイントです。これは、インテル
によるビジネスの成果や、優先度の高い
活動への取り組みをたたえるための
ものです。
インテル® パートナー・アライアンス
では、ポイントを獲得する方法が

クラウド・インサイダー・ コミュニティ



インテル® クラウド・インサイ
ダー・
コミュニティは、継続的に更新さ
れる世界水準のクラウドコンテンツ
と
ツールを提供します。メンバーは、
仲間やエコシステムとつながり、革
新的な
共同クラウド・ソリューションを市

インダストリー・インサ



ゴールドメンバーとチタンメンバーは、
特別に用意された四半期毎の
インダストリー・インサイトに
アクセスして、成長を促進できます

[詳細情報](#)

金銭的インセンティブ



メンバーになると、製品マーケティングの成功を促進する、強力な市場開発基金の活用やインセンティブ・プログラムへの参加が可能となります
インテル® パートナー・アライアンス・アクセラレーター・イニシアチブとその他の金銭的インセンティブについては、インテル担当者にお問い合わせ
ください。

リソース

インテル® パートナー・ アライアンスのカスタマーサ ポートに アクセスする方法

Intel Virtual Assistant

このチャットボットは、パートナー・アライアンスの各ウェブページの右下に設置されており、ほとんどの質問に対するセルフヘルプ、またはライブサポート・エージェントへのクイックリンクを提供します。



「サポート」ブレード

[オンライン・サポートのリクエスト](#)を送信します。

このリンクは、パートナー・アライアンスのウェブサイトでは、多くのページのフッターに表示されています。

Get Help

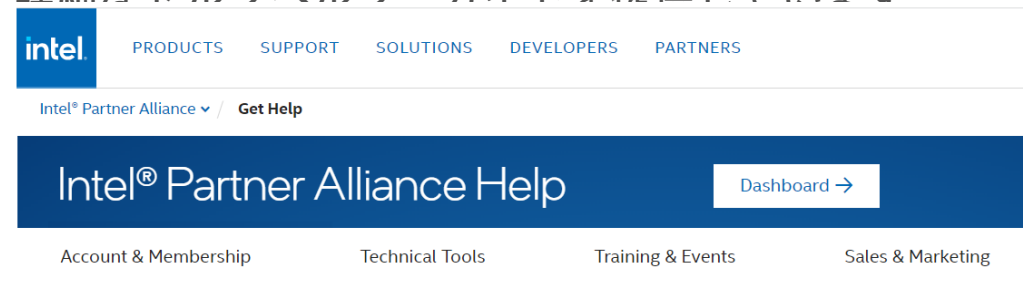
Request Support

Contact us anytime to create a support request.

[Submit request >](#)

パートナー・アライアンスの「サポート」ページ

[サポート](#)ページでは、パートナー・アライアンスのメンバーが利用できるほとんどのツールや特典に関する詳細なセルフヘルプ・ガイドを提供しています。



クラウド TV

インテル® クラウド TV は、お客様を成功に導くため、
クラウド・コンピューティングのニュース、トレンド、戦略を探ります



クラウドでの
Sapphire Rapids



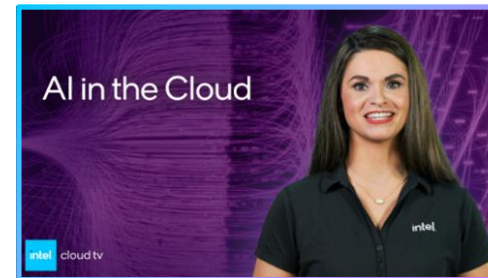
クラウドで
AI を促進



高速パスで、場所を
問わずに AI を拡張



クラウド・テクノロジー
を
用いた AI 推論



クラウドにおける AI

第4世代インテル® Xeon® スケーラブル・プロセッサー・ファミリーでの AI 情報とリソース



製品概要

[第4世代インテル® Xeon® スケーラブル・プロセッサー・ファミリー](#)

[インテル® Xeon® プロセッサー向けインテル® AI エンジンが AI パイプライン全体のパフォーマンスを向上](#)

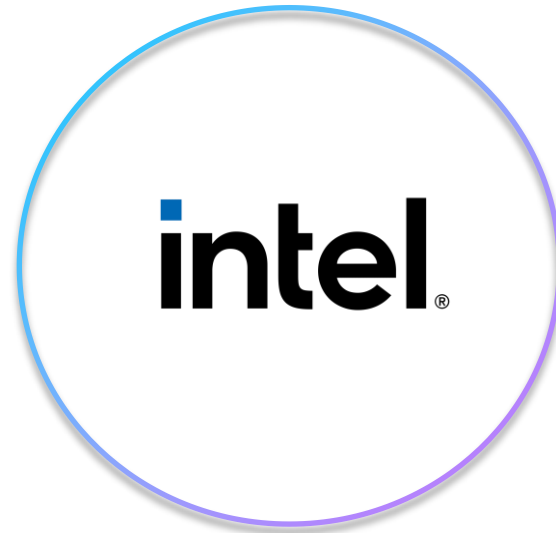


テクノロジー・ペーパー

[コンフィデンシャル・コンピューティングで AI 推論を高速化](#)

[第4世代インテル® Xeon® プロセッサーによるスケーラブルなエンドツーエンドのエンタープライズ AI](#)

[テクノロジー・イノベーターとインテルのテクノロジーによる AI イニシアチブの簡素化](#)



インフォグラフィック

[ハイパフォーマンス AI を迅速かつコスト効率よく導入](#)

[AI による迅速な ROI の向上](#)



導入事例

[富士通](#) | [Siemens](#) | [BCM](#) | [ai.io](#)



ビデオ

[インテル® AI パイプラインのビデオ](#)

[インテル® AMX: AI を高速化する大きな一歩](#)

[インテル® AI アクセラレーターのビデオ](#)

[第4世代インテル® Xeon® プロセッサー・クラウド AI ビデオ](#)



ブリーフケース

[テクノロジー・イノベーターとインテルのテクノロジーによる AI イニシアチブの簡素化](#)

関連情報



性能指標

[第4世代インテル® Xeon®
スケーラブル・プロセッサ・ファミリ](#)



カタログ

[AI 推論ソフトウェアとソリューションのカタログ](#)



追加トレーニング

[オンライン・トレーニングのスライド内リンク](#)



ビジネスレポート

[AI のハイプサイクル、2022年](#)

[デジタルファーストの経済でデジタル・
トランスフォーメーションを解き放つ：
人工知能のディスラプターになる](#)

[第4世代インテル® Xeon® スケーラブル・
プロセッサ・ファミリーによる
データセンターのパフォーマンスと
機能の加速](#)

トレーニング・アセット

AI トレーニング・アセット

人工知能 (AI)

[人工知能: 第 4 世代インテル® Xeon® プロセッサーによるワークロード・アクセラレーション](#)
すべて

[Fortanix Confidential AI を活用したオンデマンド AI ワークロード保護に関する詳細](#)
DevOps、クラウド・アーキテクチャー

[クラウドでインテル® AI を使用する理由とは？](#)
DevOps、クラウド・アーキテクチャー

[AI のクラウド展開オプション](#)
クラウド・アーキテクチャー、経営幹部

[CSP AI ポートフォリオ](#)
クラウド・アーキテクチャー、経営幹部

[データセンターからエッジまでの AI パフォーマンスを実現](#)
DevOps、クラウド・アーキテクチャー

[第 4 世代インテル® Xeon® プラットフォームのご紹介](#)
すべて

法務情報および免責事項

通知と免責事項

© Intel Corporation. Intel、インテル、Intel ロゴ、その他のインテルの名称やロゴは、Intel Corporation またはその子会社の商標です。その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

The Intel logo, consisting of the word "intel" in a lowercase, sans-serif font, is positioned in the bottom right corner of the slide. The logo is white and is set against a dark blue rectangular background.

The Intel logo is centered on a dark blue background. It features the word "intel" in a white, lowercase, sans-serif font. A small, bright blue square is positioned above the letter "i". To the right of the word "intel" is a registered trademark symbol (®). The background is a solid dark blue with several faint, semi-transparent squares of varying shades of blue scattered across it, creating a subtle geometric pattern.

intel®

バックアップ

ハードウェアとソフトウェアの複合的効果

[複合的効果のスライドはこちら](#)

システムボード	インテル® サーバー・ボード S2600STB	M50CYP2SB1U Coyote Pass	Intel Corporation / Archer City
CPU	インテル® Xeon® Platinum 8270 CPU 2.70GHz	インテル® Xeon® Platinum 8380 CPU 2.30GHz	インテル® Xeon® Platinum 8490H 1.90GHz
ソケット、物理コア / ソケット	2、26	2、40	2、60
ハイパースレディング / ターボ設定	有効 / オン	有効 / オン	有効 / オン
メモリー	12x16GB DDR4 2933MHz	16x16GB DDR4 3200MHz	16x16GB DDR5 4800MHz
OS	UB-18.04 LTS	UB-22.04 LTS	UB-22.04 LTS
カーネル	5.3.0-24-generic	5.19.0-38-generic	5.19.0-41-generic
ソフトウェア	インテル® ディストリビューションの OpenVINO™ ツールキット 2021.4	インテル® ディストリビューションの OpenVINO™ ツールキット 2022.3	インテル® ディストリビューションの OpenVINO™ ツールキット 2023.0
BIOS	SE5C620.86B.02.01. 0013.121520200651	SE5C620.86B.01.01.0006.2207150335	EGSDREL1.SYS.9409.P31.2302280828
BIOS リリース日	2020年12月15日	2022年07月15日	2023年02月28日
BIOS 設定	最適化されたデフォルト設定を選択、 保存して終了	最適化されたデフォルト設定を選択、 保存して終了	最適化されたデフォルト設定を選択、 保存して終了
テスト日	2021年6月18日	2023年06月20日	2023年05月25日
精度とバッチサイズ	int8/Batch 1	int8/Batch 1	int8/Batch 1
推論リクエストの数	52	80	120
実行ストリームの数	52	80	120
電力 (TDP) / ソケット	205W	270W	350W

ワークロード (モデル: 入力 HxW):

Inception-v4: (299x299)、Resnet-50: (224x224)、Unet-camvid-onnx-0001: (368x480)、Yolo-v3-tiny: (416x416)