

高度な分析への着手

目次

今日の IT 環境と分析について.....	2
分析の種類	2
分析ソリューション・スタック	4
4つのレイヤー.....	4
その他のテクノロジー	5
分析の開始:5つの基本ステップ.....	6
ステップ 1:解決すべきビジネス上の問題に焦点を当てます.....	6
ステップ 2:分析が組織の体質や業務をどう変えるかを理解します.....	6
ステップ 3:必要なスキルを特定し、伸ばします.....	7
ステップ 4:テクノロジーの要件について考えます.....	7
ステップ 5:データ・ソリューションを導入します.....	8
ビジネスと IT のコラボレーション.....	8
計画的な前進:デシジョンツリー.....	9
インテルだからできること.....	10
インテルIT部門の分析プラットフォーム	12
分析に関するその他の資料.....	12
インテルのプログラムや製品のウェブサイト	12
インテル IT部門ウェブサイト.....	13
インテル IT部門ホワイトペーパー.....	13
インテル・エコシステム・パートナー・ソリューションのサイト	13

今日のIT環境と分析について

ビッグデータについての今日のIT業界の反応は、過剰な期待や当初の懐疑的な見方から、どのように新たなテクノロジーを用いてデータを活用し、ビジネス価値を生み出すかという慎重な議論へと移りつつあります。もはや「ビッグデータに価値はあるのか」と問う人はいません。現時点では、人々は「どうすればデータ分析ソリューションを用いて組織に価値をもたらすことができるのか」について模索しています。

こうした変化は、すべての企業が直面している新たな現実を反映しています。すなわち、現状維持はもはや選択肢としてあり得ないということです。デジタルが加速させるこの世界で優位に立つために、組織はデータ分析ソリューションを活用して、より短時間で業界の動向をとらえ、競争力を高める必要があります。

これは抗いようのないビジネスの要求であり、そのために多くの企業が分析に投資していることは当然といえます。分析への投資は医療、小売業、製造業、金融業などあらゆる業界で行われており、大量のデータから得られる洞察が、画期的な発見、サービスの向上、顧客体験の強化など、さまざまなビジネスの目標達成につながるという事実を反映しています。

分析の種類

高度な分析を活用して洞察、革新性、競争力を得ることが組織にとっていかに貴重であるかを理解するには、現在の状況について学ぶことが役に立ちます。分析は常に進歩する科学であり、数年の間に大きく変化し、今日でも急速に進化を続けています。

現在では分析を、記述的、診断的、予測的、処方的、認知的の5つのカテゴリーに分けるのが一般的です。これらのカテゴリーは、図1の分析の成熟度モデルが示すように、階段状に積み重なっています。この段階を上げるごとに、企業はより賢く、より迅速に意思決定を行えるようになり、オンデマンド・エンタープライズに近づいていきます。現在の組織の能力および向かうべき地点を明確にするには、自分たちが現時点でこの成熟度モデルのどの段階にいるのかを理解することが重要です。

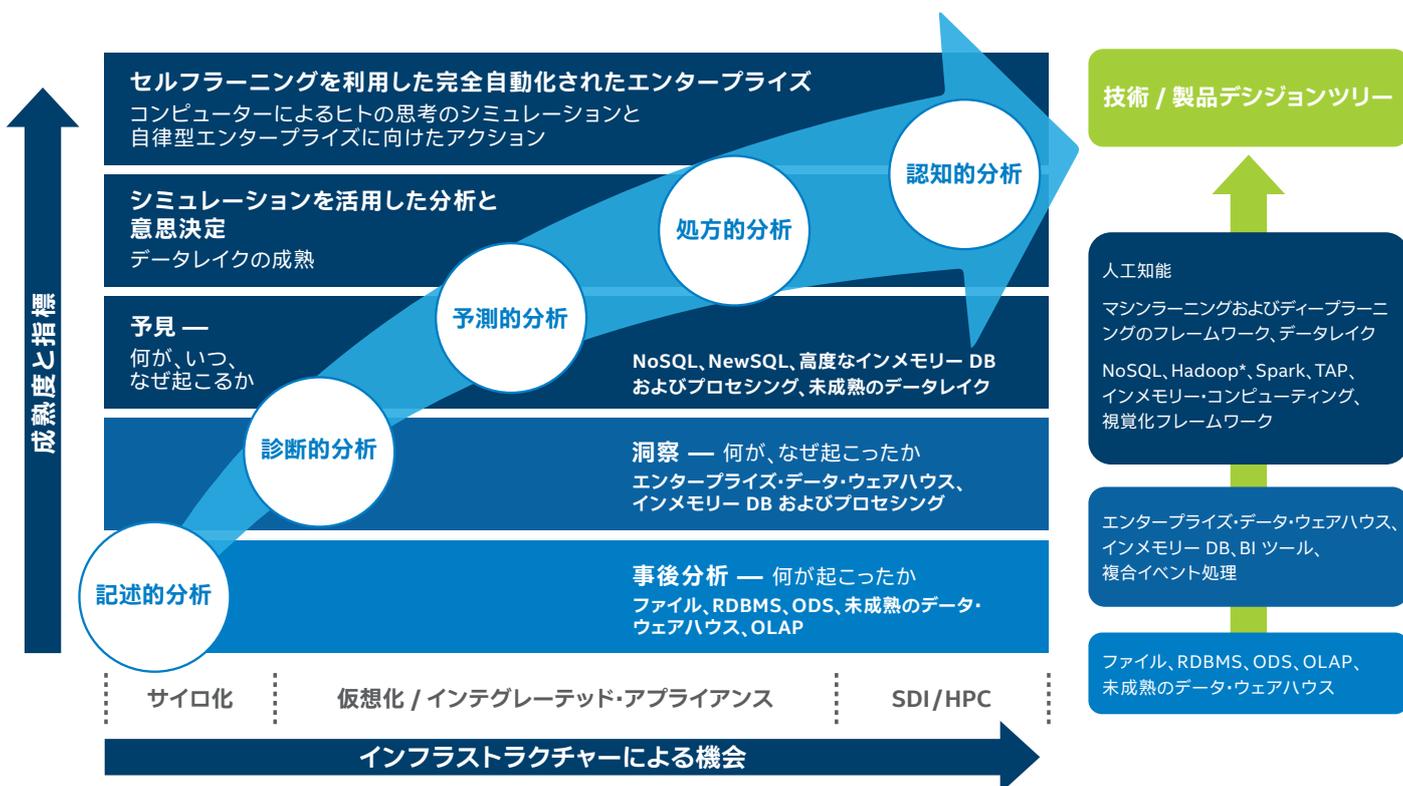


図 1. 高度な分析の成熟度:リアルタイム・エンタープライズへの移行

分析の成熟度曲線は、過去に何が起こったかを説明する基本的な分析から、意思決定を自動化する認知的分析へと進んでいきます。

分析の各カテゴリーは従来の手法と高度な手法の2つのグループに分類されます。

従来の分析

- 記述的分析は、過去に何が起こったかという問いに答えるものです。
- 診断的分析は、物事がなぜ起こったかについての洞察をもたらします。

高度な分析

- 予測的分析は、現在と過去のデータを分析することで、将来起こる可能性のある出来事についての洞察をもたらします。また、測定できる事実に基づいて、測定できない出来事を推測することを可能にします。例えば、ユーザーに空腹かどうかを尋ねることはできませんが、彼らがお昼にサンドイッチを購入しており、さらに過去にも同様の行動パターンがあった場合は、彼らが空腹であることを推測できます。
- 処方的分析は、シミュレーションやマシンラーニングなどのデータ・テクノロジーを用い、企業が望む成果を達成するためのアクションを提案します。
- 認知的分析は、マシンラーニング、ディープラーニング、ハイパフォーマンスなデータ分析などの人工知能(AI)テクノロジーを活用し、ヒトの思考に近い分析と意思決定を自動で行ったり、ヒトの意思決定を補強するようにスマートマシンを役立てたりします。

では、分析の成熟度モデルでこれらの段階を詳しく見てみましょう。今日ではほとんどの企業が、ビジネスの理解を深め、業界の動向を把握するために記述的分析もしくは診断的分析を少なからず利用しています。この経験を土台にして、企業はより高度な分析に移行することが可能となり、現在の傾向に基づいたビジネスの進路についての豊かな洞察や、最高の成果を得るためのビジネス戦略の提案を得ることができます。

組織が成長していくと、「何が、なぜ」という従来の質問から、より積極的な、将来の予測と成果に焦点が移行します。このように高いレベルの分析では、人工知能(AI)のテクノロジーが用いられます。AIを実現する主要なテクノロジーであるマシンラーニングは、特定の機能を実行するための明示的なプログラムを必要とせずに、マシンに行動や思考をさせる計算方法です。アルゴリズムのセット、言い換えるなら数学的モデルのセットが、データから「学習」するのです。これらのモデルは、特定のタスクでのパフォーマンスを経験に基づいて改善してだけでなく、新しいデータについての予測も行います。

さらに先を見通している企業は認知的分析の検討を始めています。このレベルまで成熟した分析手法では、ヒトの経験はスマートマシンで拡張され、ヒトとマシンの自然な相互作用によるコラボレーションが、ビジネスに画期的な洞察をもたらします。

IDCの予測

2020年までに、企業による分析への新規純投資の40%が、予測的分析および処方的分析に集中することになるでしょう。¹

分析ソリューション・スタック

分析ソリューション・スタックは、インフラストラクチャー、データ、分析、アプリケーションの4つのレイヤーを内包し、それらすべてをテクノロジーでつないでいます。これら4つのレイヤーにおいてテクノロジーはお互いを補完し、既存のデータ・マネジメント・アーキテクチャーの利点も活用しながら、柔軟なビッグデータ・プラットフォームとして従来の分析と高度な分析の両方を可能にします。

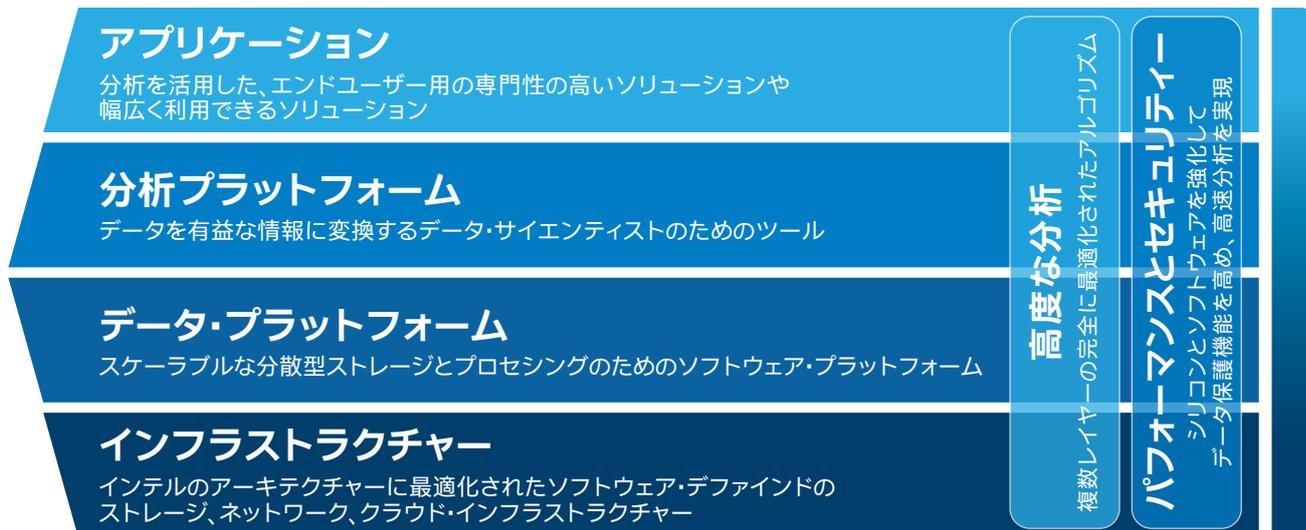


図2. 分析ソリューション

分析ソリューション・スタックは、インフラストラクチャー、データ、分析、アプリケーションの4つのレイヤーにまたがります。

4つのレイヤー

インフラストラクチャー・レイヤー

目標がリアルタイム分析のためのスケールアップ、もしくは分析環境にビッグデータのセットを含むためのスケールアウト、またはその両方だったとしても、分析ソリューション・スタックは、ハイパフォーマンスかつ拡張性の高いインフラストラクチャー・レイヤーの上に構築されます。データの取得、保管、保護、さらに商用およびオープンソースの分析ソリューションの実行はこのテクノロジーによって支えられています。ユースケースにより異なりますが、一般的にインフラストラクチャー・レイヤーは、オープンソースの分散処理フレームワーク、非リレーショナル分析データベース、分析アプリケーションを組み合わせて作動させる基盤となります。

データレイヤー

データレイヤーは、企業が分析を行うデータが格納される場所です。成熟度の低い従来の分析では、データレイヤーは主にリレーショナル・データベースで構成されていました。現在ではストリーミング・データや非構造化データを保存し、分析する必要性から、エンタープライズ・データ・ハブやデータレイクの基礎にもなり得るHadoop* 分散ファイルシステム(HDFS)がこのレイヤーを補完しています。また、NoSQLのデータベースでデータを保管することもできます。今後多くの企業が、自社のERPやIoTからのストリーミング・データなどのさまざまなデータを、SAP HANA* のようなリアルタイム分析環境に取り込むようになるでしょう。

分析レイヤー

分析レイヤーは、分析のエンドユーザー・アプリケーションにビルディング・ブロックを提供します。分析レイヤーでは、Spark*、Storm* などのオープンソース製品や、Microsoft、SAP、SAS、Oracle、IBMなどの大手企業、あるいはその他の小規模なイノベーターのスイートを使用することができます。

アプリケーション・レイヤー

ソリューション・スタックの最上位層には、すぐに導入できて、業界を問わずにさまざまなユーザータイプのデータに分析を提供する分析アプリケーションが含まれます。

その他のテクノロジー

データレイク

データレイクは、構造化データと非構造化データの両方を大量に格納できる保管庫（リポジトリ）です。データレイクはデータウェアハウスに比べて汎用性が高く、データモデルやファイル形式が異なるデータを、さまざまなストリームから、さまざまな頻度で取り込むことができます。ガートナーは次のように補足しています。「データレイクは元のデータソースに加わる、さまざまなデータ資産のストレージ・インスタンスの集合です。これらの資産は、ソースの形式に、かなり近い、もしくは正確なコピーとして格納されます。データレイクの目的は、未精製なデータを非常に高いスキルを持つアナリストだけに提示して、従来の分析データストア（例えば、データマートやデータウェアハウスなど）に存在するような情報記録システムの妥協を一切気にすることなく、自分たちのデータの精製や分析テクニックを検証できるように支援することです」²

NoSQL データベース

NoSQLのデータベースは、ウェブスケールで高い性能と可用性を持つストレージを提供します。大量のデータストリーム、フレキシブルなスキーマとデータ型を、短いレスポンス時間で処理するのに便利な非リレーショナルデータベースです。NoSQLデータベースは分散型で耐障害性に優れたアーキテクチャーを採用しているため、システムの信頼性と拡張性が確保されています。NoSQLデータベースの例として、Apache HBase、Apache Cassandra*、MarkLogic*、MongoDB*、Apache CouchDB*ソフトウェアなどが挙げられます。

これら4つのレイヤーにおいてテクノロジーはお互いを補完し、既存のデータ・マネジメント・システムの利点も活用しながら、柔軟なビッグデータ・プラットフォームとして従来の分析と高度な分析の両方を可能にします。さらに、パフォーマンスとセキュリティに関するソリューションや人工知能など、その他のテクノロジーもソリューション・スタックのすべてのレイヤーに活かされており、データを厳重に保護しながら、高度な分析から得られる洞察を強化します。

人工知能(AI)

データサイエンティスト、デベロッパー、研究者は、これまでには望むこともできなかった洞察を得るために、アルゴリズムもしくは数学的モデルがデータから「学習」するマシンラーニングなどの、AIの前段階であるツールを活用しています。分析ソリューション・スタックは、複数レイヤーの完全に最適化されたアルゴリズムを用いてマシンラーニングを促進し、明白な指示を与えることなく、コンピューター・モデルによって複雑な非構造化データから洞察を得ることができます。

ディープラーニングは、急速に成長しているマシンラーニングの1つの部門です。これは、複数レイヤーのディープグラフを用いて経験から学習する人工ニューラル・ネットワークを使用したもので、新しいデータからの洞察を推測することができるモデルを形成します。このアプローチにより、画像認識、音声認識、自然言語処理など、マシンよりも人間の方が有利とされてきた多くの分野でブレークスルーがもたらされました。飛躍的に向上するマシンの性能とこれらのアプローチの組み合わせにより、今日大きなビジネス価値が生まれており、あらゆる業界が変革の時を迎えています。

パフォーマンスとセキュリティ

パフォーマンスとセキュリティは、高い競争力を持つオンデマンドかつリアルタイムな企業にとって最も重要な項目です。だからこそ、分析ソリューション・スタックでは、さまざまな観点からシリコンとソフトウェアを強化することでデータ保護機能を高め、高速分析を実現しています。

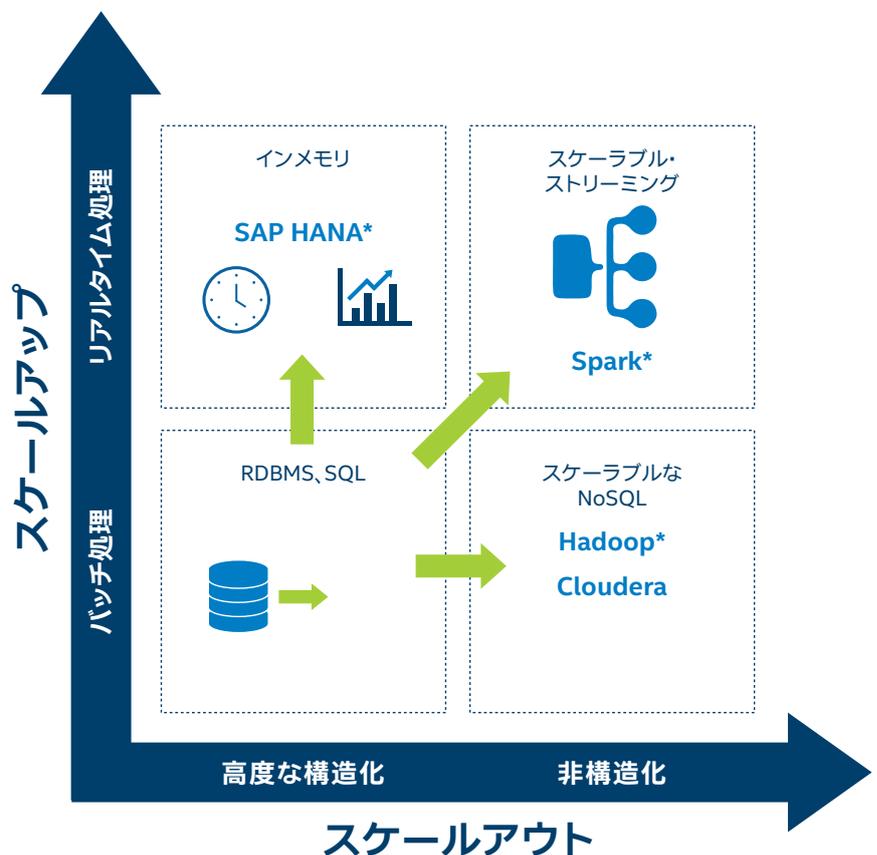


図2. 分析ソリューションは、スケールアウトやスケールアップによってさまざまな種類、ボリューム、速度のデータに対応することができます。

NoSQL データベース

NoSQLのデータベースは、ウェブスケールで高い性能と可用性を持つストレージを提供します。大量のデータストリーム、フレキシブルなスキーマとデータ型を、短いレスポンス時間で処理するのに便利な非リレーショナルデータベースです。NoSQL データベースは分散型で耐障害性に優れたアーキテクチャーを採用しているため、システムの信頼性と拡張性が確保されています。NoSQL データベースの例として、Apache HBase、Apache Cassandra*、MarkLogic*、MongoDB*、Apache CouchDB*ソフトウェアなどが挙げられます。

分析の開始: 5つの基本ステップ

ここまで、ビッグデータを取り巻くIT環境、ビッグデータが組織にもたらす潜在的価値、そして構造化、半構造化、非構造化のどのタイプのデータリソースからも洞察を得ることができるテクノロジーについて説明してきました。さらに、適切なインフラストラクチャーを導入するための基本と、ビッグデータ・イニシアチブを支えるスムーズな稼働についての概要にも触れました。

これから説明するのは5つの基本ステップであり、この手順に従うことで、ビッグデータ分析プロジェクトを開始できます。

ステップ1: 解決すべきビジネス上の問題に焦点を当てます。

ビジネスユーザーと協力して、ビジネスチャンスを明確にしましょう。

- 分析、データ、コンプライアンスの各担当者、データサイエンティスト、データスチュワード、シチズン・データサイエンティスト、そしてデベロッパーを含むビジネスユーザーを特定し、彼らと協力して組織におけるビッグデータ分析の最大のビジネスチャンスを特定します。
- 現在のビジネスの課題、特に、解決が難しくコストがかかるもの、現在のデータソースと分析システムでは解決不可能なものについて検討します。または、データソースが新しく、構造化されていないために、今日まで着手できていない問題について検討します。
- ビジネスチャンスをリストにしたら、優先順位をつけ、投資収益率 (ROI) をはっきりと測定できるプロジェクトを選択します。最適なプロジェクトを決定するために、次の質問に対する答えを検討の基準としてください。
 - 達成しようとしていることは何ですか？
 - このプロジェクトは組織の戦略的目標と一致しているか？
 - プロジェクトに対してマネジメントのサポートは得られるか？
 - 従来の分析では得られなかったような洞察がビッグデータ分析に期待できますか？
 - プロジェクトの結果を参考にしてどのようなアクションを取れますか？
 - 投資に対する予想利益はどれくらいですか？
 - 半年から1年ほどで価値を生み出せるプロジェクトですか？
 - 必要としているデータが手に入りますか？すでに持っているものは何ですか？購入する必要があるものは何ですか？
 - データはリアルタイムのものでしょうか、それとも過去のものでしょうか？

ビッグデータの戦略とアプローチを考える際は会社の経営層に協力を求めましょう。以下について明確にしてください：

- 分析のためのビジネスケース - 分析がどのようにビジネスに価値をもたらすかを定義します。分析ソリューションで解決を図るビジネスの主な課題を特定します。
- 短期、中期、長期の目標 - 分析の目標を達成するために主要なフェーズを定義します。

ステップ2: 分析が組織の体質や業務をどう変えるかを理解します。

データ分析ソリューションについての理解を深めましょう。

- IT部門やビジネス部門の同僚と話をします。
- テクノロジーのスピードに取り残されないよう、分析に関するインテルのリソースを活用してください。
- ベンダーの提案内容を理解します。
- Apache Hadoop* プロジェクトが提供するチュートリアルなどを活用し、ユーザー用の資料にも目を通します。

インフラストラクチャーおよび運用上のニーズを評価しましょう。以下について考察してください：

- 組織のITインフラストラクチャーの現状と未来 – 現在のデータセンターはビッグデータプラットフォームに対応できますか？現在使用しているデータセンターのテクノロジーを確認し、必要であれば、コンピューティング、ストレージ、ネットワークなどのリソースのアップグレードを検討します。
- データソースとデータ品質 – 内部のプライマリー・データソースは何ですか？外部から入手する可能性のあるほかのデータは何ですか？データの品質をどのように確保しますか？
- 分析のプラットフォームとツール – ソリューションの構築に使用するプラットフォームは何ですか？目的達成のために必要なソフトウェアとツールは何ですか？
- 成功の測定基準 – システムのパフォーマンスをどのように測定しますか？どれくらいのジョブが投入され、並行処理され、効率的に完了できるかを、成功の基準とします。

ステップ3: 必要なスキルを特定し、伸ばします。

ビジネスとITのそれぞれに必要なスキルについて理解し、それらのスキルを得る計画を立てましょう。

- 分析のイニシアチブを効果的に達成するために必要なスキルを特定します。
- 必要なリソースが社内にあるかを確認します。
- 会社の中でスキルを構築できるかを確認します。
- ビジネス部門やIT部門内で分析の専門家がどこに属するべきかを決定します。
- 必要に応じて、スキルを持つ人材を新しく雇うか、特定の業務を外部委託します。

ステップ4: テクノロジーの要件について考えます。

現在の能力と、将来的に目指す能力とのギャップを認識しましょう。

- データを収集、浄化、統合して使用可能なフォーマットにするために、データ品質に関するどのような要件が追加が必要ですか？
- データの分類、関連性の定義、保管、分析、アクセスのそれぞれで、どのようなデータ管理ポリシーを策定する必要がありますか？
- 拡張性、低レイテンシー、パフォーマンス(コンピューティング、ストレージ、ネットワーク機能など)を確保するために、インフラストラクチャーにはどのような能力が求められますか？
- 種類が異なる大量のデータを高速で検索するために、NoSQLデータベースなどの専門的なコンポーネントを追加する必要がありますか？
- リアルタイム・データのストリームを安定して処理したい場合は、どんなインフラストラクチャーとメモリー機能を追加する必要がありますか？MPPインメモリー分析アプリケーションが必要ですか？CEPソリューションが必要ですか？
- 目的とするアウトプットを生成するために必要な分析クエリーと分析アルゴリズムを特定します。
- デリバリーモデルとしてクラウド・コンピューティングを検討している場合は、プライベート、パブリック、ハイブリッドの中でどのクラウド環境が適していますか？
- データはユーザーにどのように提示されますか？得られた結果は、上級エグゼクティブや情報専門家を含む幅広いビジネスユーザーにとって簡単に理解できる形で示す必要があります。

ビジネスとITのコラボレーション

企業の行く末を大きく左右する分析への取り組みは、周囲の協力なくしては達成できません。ビジネスリーダーとの強力なパートナーシップの下、IT部門はビッグデータの可能性を見出し、必要なサポートを受けながら前進していく必要があります。また、分析を活用するには、複雑なビジネス上の問題を的確に把握し、洞察を発見し、システムを統合し、巨大なデータベースを構築し、分散ソフトウェア・フレームワークを管理、運営していくために、経営、技術、分析のそれぞれにおいて新しいスキルが必要になります。

ビジネスに高度な分析を導入するということは、将来的には、データが経営判断における最も大事な位置を占めるようになることを意味します。そのため、データとそれを扱う各プロセスに適切な責任者を用意する必要があります。多種多様で膨大な量のデータを管理、説明、分析するための新しい方法に対処するには、組織内のデータに関連する役割を再考しなければなりません。

データに関する組織内での役割には、すでに確立されているものもあれば、比較的目新しいものもあります。例えば、チーフ・データサイエンティスト、チーフ・アナリティクス・オフィサー、シチズン・データサイエンティストなどは、最近になって登場した新しい役割です。

- チーフ・データサイエンティストは、コンピューターサイエンス、数学、統計モデリング、分析の専門知識を兼ね備えており、ビッグデータから実用的な洞察を引き出します。
- チーフ・アナリティクス・オフィサーは、データから洞察と価値を得るために、組織での分析ツールの使用を監督します。
- シチズン・データサイエンティストは、データから洞察を得るために分析ツールを使用しますが、ビッグデータや分析を主業務とするとはありません。この役割を担う人材は、組織内のどの部門に属していてもかまいません。

注目すべき傾向をもう1点挙げるとすれば、データ分析チームは、データを活用したビジネス展開をサポートするために、異なる部門の人々と密接に連携する機会が多くなるということです。

ステップ5: データソリューションを導入します。

プロジェクトのユースケースを発展させましょう。

- プロジェクトの遂行に必要なユースケースを特定します。
- データの流れを図にすると、必要な分析能力を定義するのに役立ちます。
- 使用するデータと除外するデータを選り分けます。意義のある洞察に結びつく戦略的データだけに絞りましょう。
- 各データの相互関係とビジネスルールの複雑さを判断します。
- インタラクティブクエリーや予測的分析といった高度な分析、もしくはリアルタイム・データストリームに対応する必要があるかを検討します。

実機バージョンのテスト環境を整備しましょう。

- リファレンス・アーキテクチャーを自分たちの組織に適応させます。インテルは一流のパートナーと連携し、ビッグデータのユースケースに関連したインテル® クラウド・ビルダーズのプログラムの一環として、役に立つリファレンス・アーキテクチャーを開発しています。
- プレゼンテーションレイヤー、分析アプリケーションレイヤー、データウェアハウスを定義し、必要に応じて、プライベート、もしくはパブリックベースのクラウド・データ・マネジメントも定義します。
- ユーザーが結果を有意義な形でプレゼンテーションするためのツールを決定します。ツールがユーザーに普及するかどうかは、プロジェクト全体の成功を大きく左右します。

計画的な前進: デジジョンツリー

図4に示すようなデジジョンツリーをたどることで、組織にとって最適な分析ソリューションを判断する材料を得ることができます。デジジョンツリーの質問に順次答えていくことで、自分が求めている特定の要件やテクノロジーのニーズを、チーム全体として理解しやすくなります。

図4のデジジョンツリーでは、テクノロジーと分析に関する要素が青の背景で表されており、組織に関する要素が緑の背景で表されています。また、LOBおよびテクノロジーの関係者によって確実に把握されている明確なユースケースから開始します。このユースケースは、上の図1で紹介した分析の成熟曲線上の組織の位置を反映していなければなりません。例えば、成熟度モデルにおいて診断的分析のレベルにあり、日次のバッチ処理でデータを取得している組織では、リアルタイムの予測を必要とするユースケースに対応する準備ができていないとは言えません。

対象となるユースケースをよく検討し、それがリアルタイム分析を必要としているのか、またはバッチで十分なのかを判断する必要があります。また、ユースケースに必要なデータも定義しなくてはなりません。内部のERPシステムから抽出する構造化データが必要なのか、または外部から、場合によっては非構造化データを取り入れる必要があるのか、などといった観点から検討しましょう。図4のフローチャートは、組織やテクノロジーに関する要因が、組織が必要とするテクノロジーやスキルにどのように影響するかを示す、優れたガイダンスとしても有用です。

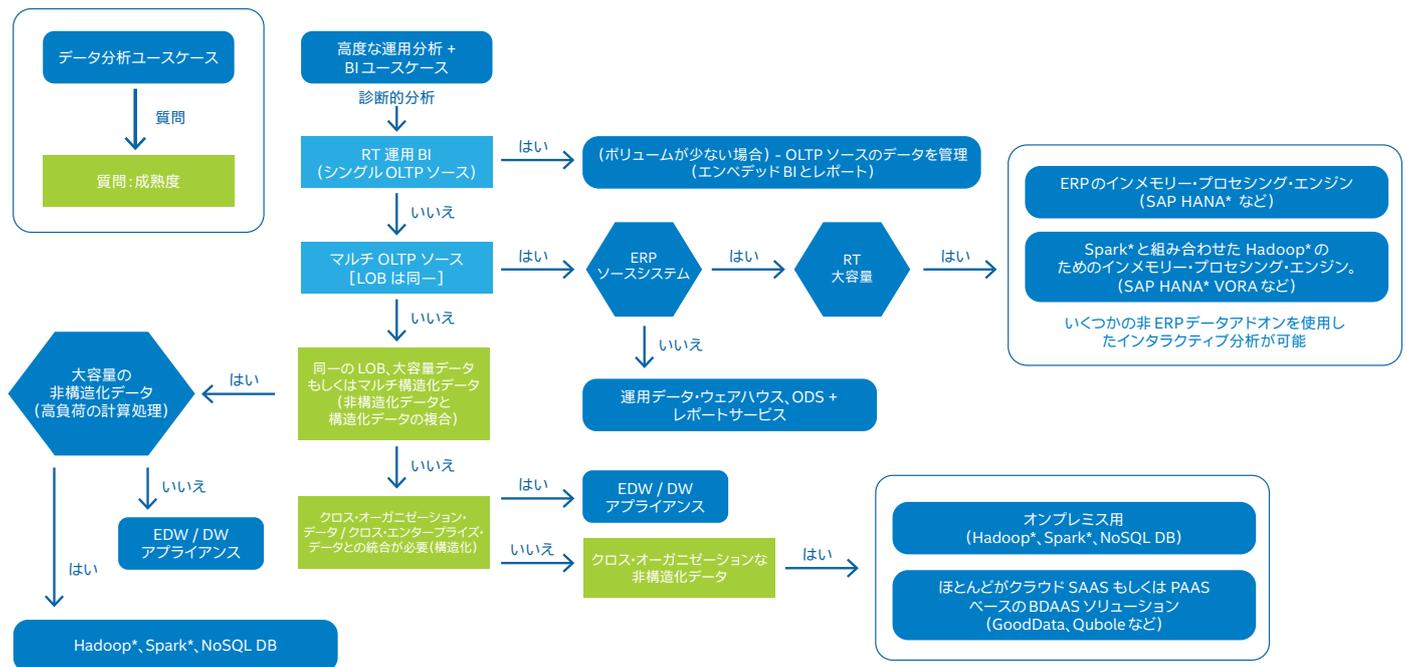


図 4. 分析のためのデジジョンツリー

インテルだからできること

ビッグデータ・インフラストラクチャーの提供者であるだけでなく、オープンソース・プロジェクトに多大な貢献を行い、ISVの最適化に積極的なサポートを行ってきたインテルは、データ分析ソリューション・スタックのほぼすべてのレベルに関わっており、革新的な分析アプリケーション、相互運用性、インテル® アーキテクチャー上での優れたパフォーマンスを実現しています。インフラストラクチャー、データ、分析、アプリケーションの各レイヤーを包括する高度な分析ソリューション・スタックをテクノロジーで支えているのが、インテルとそのエコシステム・パートナーなのです。

また、インテルは先進的なコンピューティング、ストレージ、ネットワークを通じてソリューション・スタックの基盤を構成するだけでなく、インフラストラクチャーの上に構築されるレイヤーで使用されるソフトウェア・ライブラリーやアクセラレーターなど、セキュリティとパフォーマンスを強化するソリューションも提供しています。



図5. 分析ソリューション・スタック

インテルのテクノロジーは、分析ソリューション・スタックの全体にわたって貢献しています。

ここでは、データセンター・インフラストラクチャー全体にわたって分析のための基盤を提供するインテルの製品やテクノロジーのいくつかを紹介します。

コンピューティング

インテルの豊富なプロセッサ製品はワークロードに最適化されており、分析に関するあらゆるニーズに応えることができます：

- インテル® Xeon® プロセッサ E5 ファミリーは、分析のスケールアウト・デプロイメントや、推論としても知られるマシンラーニングの実行を可能にする機能に、パフォーマンスと効率性を理想的なかたちで組み合わせています。
- インテル® Xeon® プロセッサ E7 ファミリーは、インメモリー分析とスケールアップ・アーキテクチャーを可能にする拡張メモリー容量を提供します。
- インテル® Xeon Phi™ 製品ファミリーは、シミュレーションやマシン・ラーニング・アルゴリズムの開発などのタスクにおいて優れた働きをします。

ストレージ

データストレージの容量はますます増加していますが、一方でフロアスペース、電力、予算はなかなか増やすことができません。ソリッドステート・ドライブなどに代表される、従来のストレージメディアの常識を変える発明により、ストレージのコスト、容量、レイテンシーが再定義されつつあります。インテル® Xeon® プロセッサと同時に設計されたストレージ・ソリューションは、特にソフトウェア・デファインド・インフラストラクチャーを主眼において設計および最適化された、高速かつ信頼性の高いストレージを実現します。

今日、インテルは次世代のストレージ・ソリューションの展開に力を入れており、インテルのプロセッサ、チップセット、ファームウェア、ソフトウェア、ドライバーによる機能の強化でシームレスなパフォーマンスを発揮するインテル® Solid-State Driveのような製品や、驚くほど高速でありながら低価格で、高い耐久性を実現した新しい3D XPoint™ テクノロジーなどを提供しています。

ネットワーク

コンピューティング・インフラストラクチャーにデータレイクやデータ・ウェアハウスからデータを移動する際の遅延は、リアルタイムの運用においてコスト要因となります。データ移動の際の遅延を最小化するため、インテルはこれまで以上の高速化を可能にするファブリック、インテル® Omni-Pathアーキテクチャーを用意しています。リモート・ダイレクト・メモリー・アクセス (RDMA) を支えとするインテル® Omni-Pathアーキテクチャーは、ハイパフォーマンスなデータ分析とコンピューティングのための次世代ファブリックです。また、インテルはデータセンター用のイーサネット・アダプターとコンバージド・アダプターの市場において世界的なシェアを持っており、現在までに、実に10億個以上のイーサネット・ポートを世に送り出してきました。

ハイパフォーマンス・コンピューティング

ますます複雑化する課題に対処するために、多くの企業が分析に着目しています。求められているのは、ほぼリアルタイムで結果を得ることであり、そのためにHPCへの関心が高まっています。HPCクラスターを導入する際の煩雑な手間を削減できるように、インテル® スケーラブル・システム・フレームワーク (インテル® SSF) が用意されています。これにより、共通のインフラストラクチャー上で分析ワークロードを実行するための画期的なパフォーマンスが実現し、イノベーションを加速させることができます。

セキュリティ

インテルのハードウェアおよびソフトウェアのセキュリティ・ツールは、既知の脅威の遮断、セキュリティ侵害の特定、セキュリティへの迅速な修正を行うことで、ネットワークを安全に保護します。インテル® AES New Instructions (インテル® AES-NI) などのセキュリティ機能により、暗号化のパフォーマンスが向上し、より多くのデータを安全に保管、転送、分析できるようになります。

ソフトウェア・ライブラリーおよびアクセラレーター

これまでに述べてきたタスクに関して、いずれの成功の鍵もほとんどは計算速度にあります。これを実現するのが、インテル® アーキテクチャー上でより高いパフォーマンスを発揮するためにフレームワークを最適化する、革新的なソフトウェア・ライブラリーとアクセラレーターです。

- **インテル® データ・アナリティクス・アクセラレーション・ライブラリー** (インテル® DAAL) は、ソフトウェア・デベロッパーがアプリケーションの開発にかかる時間を短縮し、より高いパフォーマンスを発揮できるように設計されています。インテル® DAAL は、利用可能なコンピューティング・リソースを用いて、アプリケーションが、より多くのデータセットを高速に分析し、これまで以上に信頼できる予測を立てられるようにサポートします。
- **インテル® マス・カーネル・ライブラリー** (インテル® MKL) は、数学処理ルーチンを加速することで、アプリケーションのパフォーマンスを向上させ、開発時間を短縮します。インテル® MKL は、高度にベクトル化およびスレッド化された線形代数、高速フーリエ変換 (FFT)、ベクトル演算、統計機能に対応しています。これらの処理能力のすべてを活用する最も簡単な方法は、丁寧に最適化された数学ライブラリーを使用することです。たとえ最高のコンパイラーでも、手動で最適化されたライブラリーが引き出すパフォーマンス・レベルには敵いません。
- **ディープ・ニューラル・ネットワーク向けインテル® マス・カーネル・ライブラリー** (インテル® MKL-DNN) は、ディープ・ニューラル・ネットワーク (DNN) アプリケーションのためのDNNパフォーマンス・プリミティブを含めることにより、CPUに最適化された人工知能の運用をサポートしています。これらのオープンソース・ソフトウェア・ライブラリーは、インテル® アーキテクチャー用に最適化されたディープラーニング・アプリケーションの開発と実行をサポートします。

例えば、インテル® MKL-DNNを使用してAIのフレームワークを自身で開発したり、高度なニューラル・ネットワークを構築したりするなど、インテルのオープンソース・ソフトウェアはAIをますます身近なものにしています。これにより、ビッグデータとクラウド・アプリケーションの間のギャップを埋めることが可能になります。

インテルIT部門の分析プラットフォーム

エコシステム・パートナーとの連携による、お客様の分析プラットフォーム構築のサポートに加え、さまざまな分析ユースケースをインハウスでサポートできるように、インテルはプラットフォーム・開発におけるその専門性を注ぎ込んでいます。インテルIT部門によるBIおよび分析ツールの使用により、インテルの収益は2014年だけで3億5,100万米ドルも増加しました。³

インテルIT部門がデータ分析を活用する方法を検証しましょう。インテルIT部門では、データ・レイク・モデルを使ってビジネス・インテリジェンス (BI) データを統合および接続する分析プラットフォームを構築しています。これによって、貴重な洞察を得られるまでの時間を24時間まで短縮することが可能になります。このインテグレートド分析ハブ (IAH) が、インテルのセールスやマーケティング部門に、データに基づいたよりスマートな意思決定を行わせているのです。

その他の利点:

- データ・レイク・モデルにより、データ・サイエンティスト、アナリスト、データ・スチュワード、エンドユーザーは、より多くのデータに、より迅速かつ柔軟にアクセスできます。複数のフォーマットで、ローデータ、浄化済み、適合済みの3つの状態のデータにアクセスできるので、さまざまなレベルの分析に対応できます。
- すべてのユーザーはセルフサービスBIの作業を簡単に行えるようになります。各ユーザーは、自身が好むフロントエンド・ツールを使って、3つの状態のどのデータに対しても分析を行うことができます。
- データセットの相互接続、セルフサービスBIポータル上での視覚化、さらにレポートやダッシュボードなどの共有により、IT部門での手作業の発生を回避し、時間を短縮できます。

IAHを使用することで、データ分析のプロジェクトはマーケティング・デジタル・メディア関連の支出を四半期ごとに約17万米ドルも削減できると推定されています。

詳しくは、ホワイトペーパーをご覧ください。「[インテルの統合分析プラットフォームによる販売およびマーケティングの支援](#)」

分析に関するその他の資料

これまで紹介してきたリソースに加え、分析についてさらに理解を深めるには、以下のコンテンツも利用できます。

インテルのプログラムや製品のウェブサイト

- [インテル製品を使った高度な分析](#)
- [分析に関するブログや洞察 \(英語\)](#)
- [マシンラーニング](#)
- [人工知能](#)
- [インテル® Xeon® プロセッサ E5ファミリー](#)
- [インテル® Xeon® プロセッサ E7ファミリー](#)
- [インテル® Xeon Phi™ 製品ファミリー](#)
- [インテル® Nervana™ システム \(英語\)](#)
- [Saffron Technology \(英語\)](#)

インテルIT部門 ウェブサイト

ITベスト・プラクティスの共有

分析、ビッグデータ、IoTなどを含むITのベスト・プラクティスに焦点を当てたホワイトペーパー、レポート、その他のリソースにアクセスできます。

インテルIT部門 ホワイトペーパー

インメモリー・データ・プラットフォームによるインテルのサプライチェーンの最適化

リアルタイムでビジネスの予測的分析を行い、より優れた意思決定を短時間でできるようにする新しいインメモリー・データ・プラットフォームが、どのようにインテルのサプライチェーンに変革をもたらすかをご覧ください。

Big Data: Securing Intel IT's Apache Hadoop* Platform(英語)

Apache Hadoop*のプラットフォームをペリメーター、アクセス、可視性、データのレベルごとに保護することを目的としたインテルIT部門によるApache Sentry*とCloudera Navigator*の活用についてご紹介します。

How Intel IT Successfully Migrated to Cloudera Apache Hadoop*(英語)

インテルIT部門のApache Hadoopソフトウェアへの移行と、インテルITマイグレーションチームが作成した6つのベストプラクティスについて説明しています。

Improving Manufacturing with Advanced Data Analytics(英語)

インテルIT部門が、センサー、IoT、データ分析の活用によって、いかに製品品質の向上、コスト削減、市場投入までの時間短縮を実現しているかをご確認ください。

Using Big Data in Manufacturing at Intel's Smart Factories(英語)

リアルタイムのデータによる自動コントロール・システムを実現するために、インテルの先進的な工場が、エッジ・コンピューティングやIoTをいかに活用しているかをご確認ください。

インテル・エコシステム・パートナー・ソリューションのサイト

インテル®アーキテクチャー上での相互運用性とパフォーマンスの向上を目指してソリューションを最適化するために、インテルはエコシステム・パートナーと積極的な連携を行っています。下記のパートナーページではこういったコラボレーションの例についてご紹介しています。

Cloudera

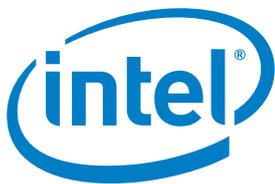
インテルはClouderaとのパートナーシップにより、セキュリティー、パフォーマンス、マネジメント、ガバナンスについて、Hadoop*フレームワークにエンタープライズ・クラスのイノベーションをもたらすことに寄与しました。

SAS

SAS* Analytics 9.4はインテル® Xeon® プロセッサー E7ファミリー向けに最適化されているため、単一の強力な新システムで予測的分析ワークロードのパフォーマンス向上を確認できます。

SAP

SAP HANA*とインテル® Xeon® プロセッサー E7ファミリー(インテル® TSX対応)の相互最適化により、リアルタイムのトランザクションおよび分析でパフォーマンスが向上します。



¹ IDC, Big Data Forecast, November 2015.

² Gartner IT Glossary: Data Lake.

³ ITイノベーションでインテルの変革を実現: 2014~2015年インテルITビジネスレビュー(年刊版)

インテル® テクノロジーの機能と利点はシステム構成によって異なり、対応するハードウェアやソフトウェア、またはサービスの有効化が必要となる場合があります。実際の性能はシステム構成によって異なります。絶対的なセキュリティを提供できるコンピューターシステムはありません。詳細については、各システムメーカーまたはお近くの販売店にお問い合わせいただくか、<https://www.intel.co.jp/content/www/jp/ja/big-data/sap-hana-real-time-analytics-solution-brief.html> (英語) を参照してください。

Intel、インテル、Intel ロゴは、アメリカ合衆国および/またはその他の国における 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。

インテル株式会社

〒100-0005 東京都千代田区丸の内 3-1-1

<http://www.intel.co.jp/>

© 2017 Intel Corporation. 無断での引用、転載を禁じます。

2017年8月

336403-001JA

JPN/1708/PDF/WLT/MKTG/YA