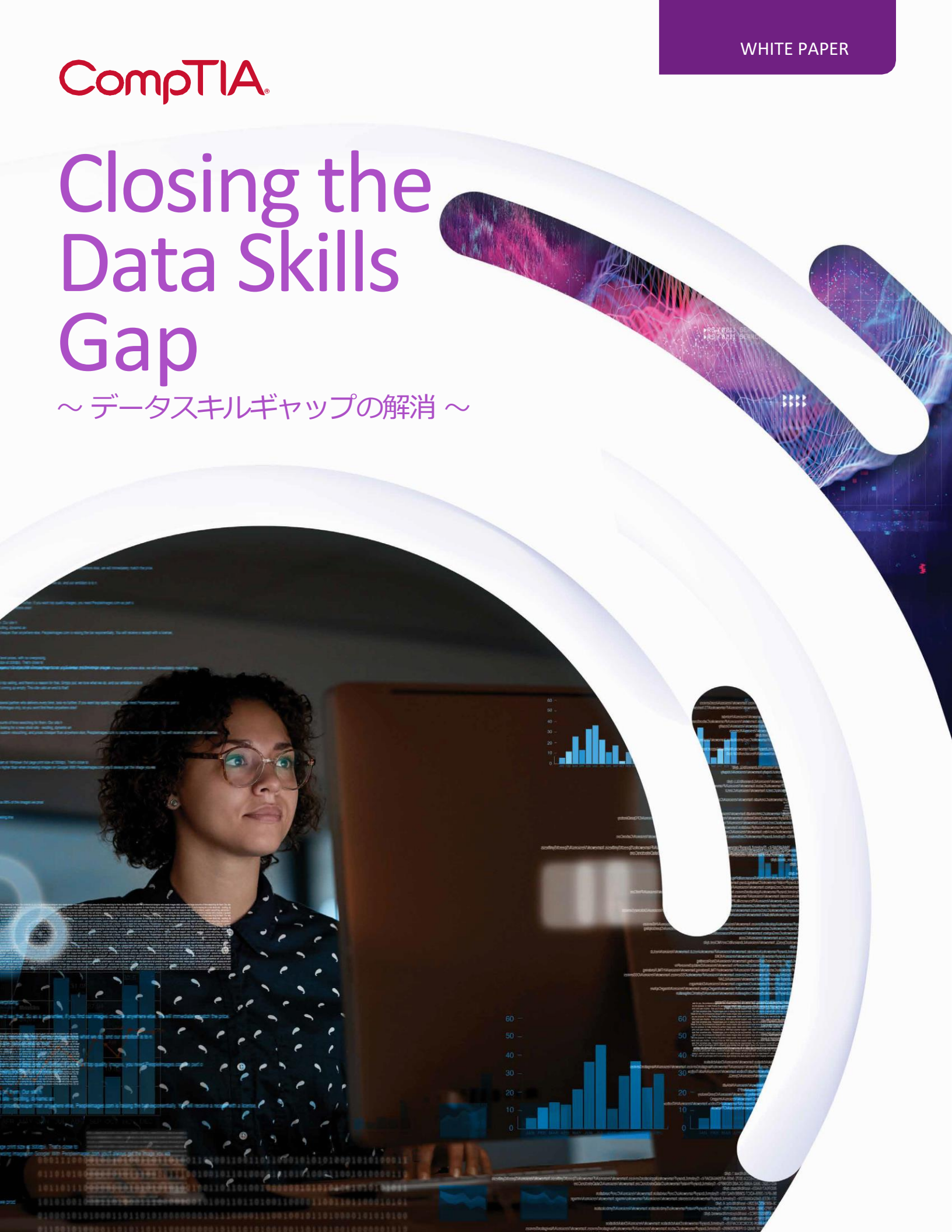


Closing the Data Skills Gap

～ データスキルギャップの解消 ～

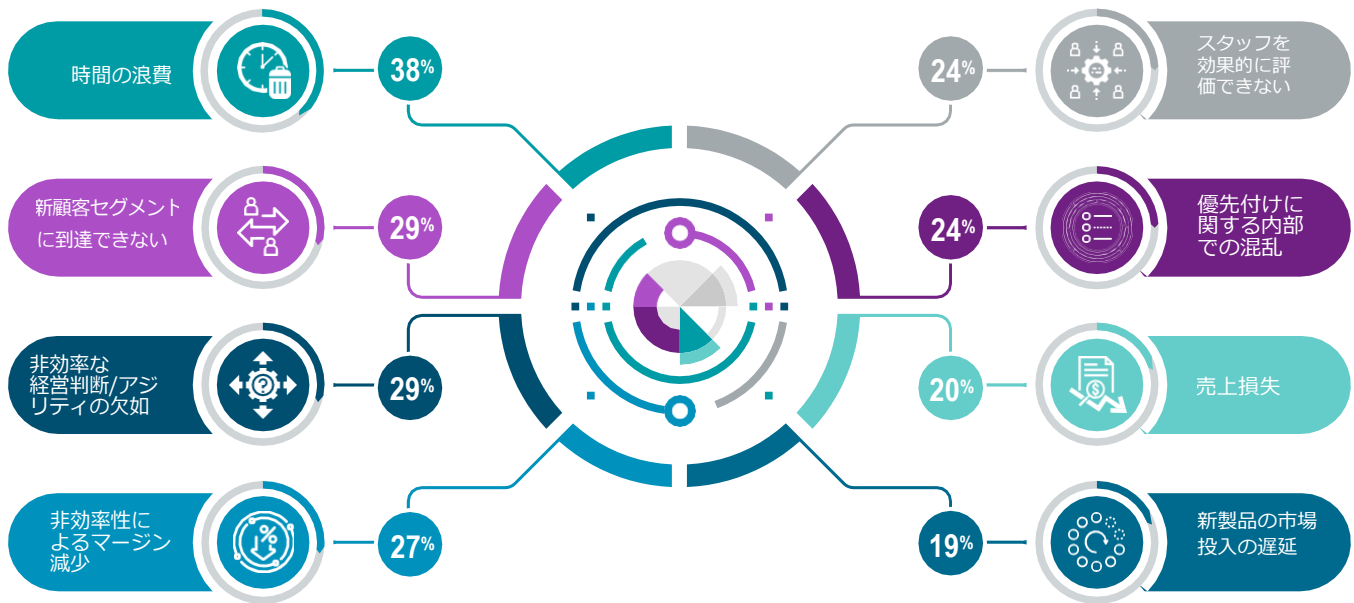


デジタル経済におけるデータの重要性については、多くのことが語られてきました。デジタルデータは何十年も前から存在していますが、人工知能（AI）、予測分析、カスタマーパーソナライゼーション（顧客の個別化）の時代において優先事項とされています。データは、ビジネスの「新しい通貨」や「新しい石油」と呼ばれ、これは業務構造全体を支えるほどの大きな資産です。しかし、データ価値が認められているにもかかわらず、企業はデータの規律やスキルを構築する初期段階にあります。

今日のビジネスにとってデータがそれほど重要なのはなぜなのでしょう？簡単に言えば、データ管理が不十分であったり、データ分析が不十分であったりすると、収益に影響が出てしまうということです。組織はデータを収集してきましたが、包括的な運用アプローチがなく、必要なスキルの確保ができていません。インテリジェンスを伴わないデータ氾濫とも言えます。ほとんどの企業はすべての部門にデータサイロがあり、企業データの全体像を把握することができません。データは、ソーシャルメディアやスマートデバイスなどの新しいソースからストリーミングされていますが、このデータを経営判断に結び付けるための仕組みがありません。歴史的な記録が豊富にあるものにあ、これらを将来プランのため使うための専門知識が明らかに不足しています。

企業は、うまくいっていないデータ管理や不十分なデータ分析の影響を感じ始めています。その最たるものが時間の浪費です。組織の様々な場所で必要データを探す必要性から、イノベーションやコア業務のために使うことができる時間が失われてしまうのです。テクノロジーの共通目的でもある効率性の向上は、適切に設計されたシステムとワークフローにより得られるものです。データ管理と分析は、特定の目的のためのツールではなく、包括的なプログラムとして扱うべきなのです。

データ活用がうまくいかないことで生じる影響

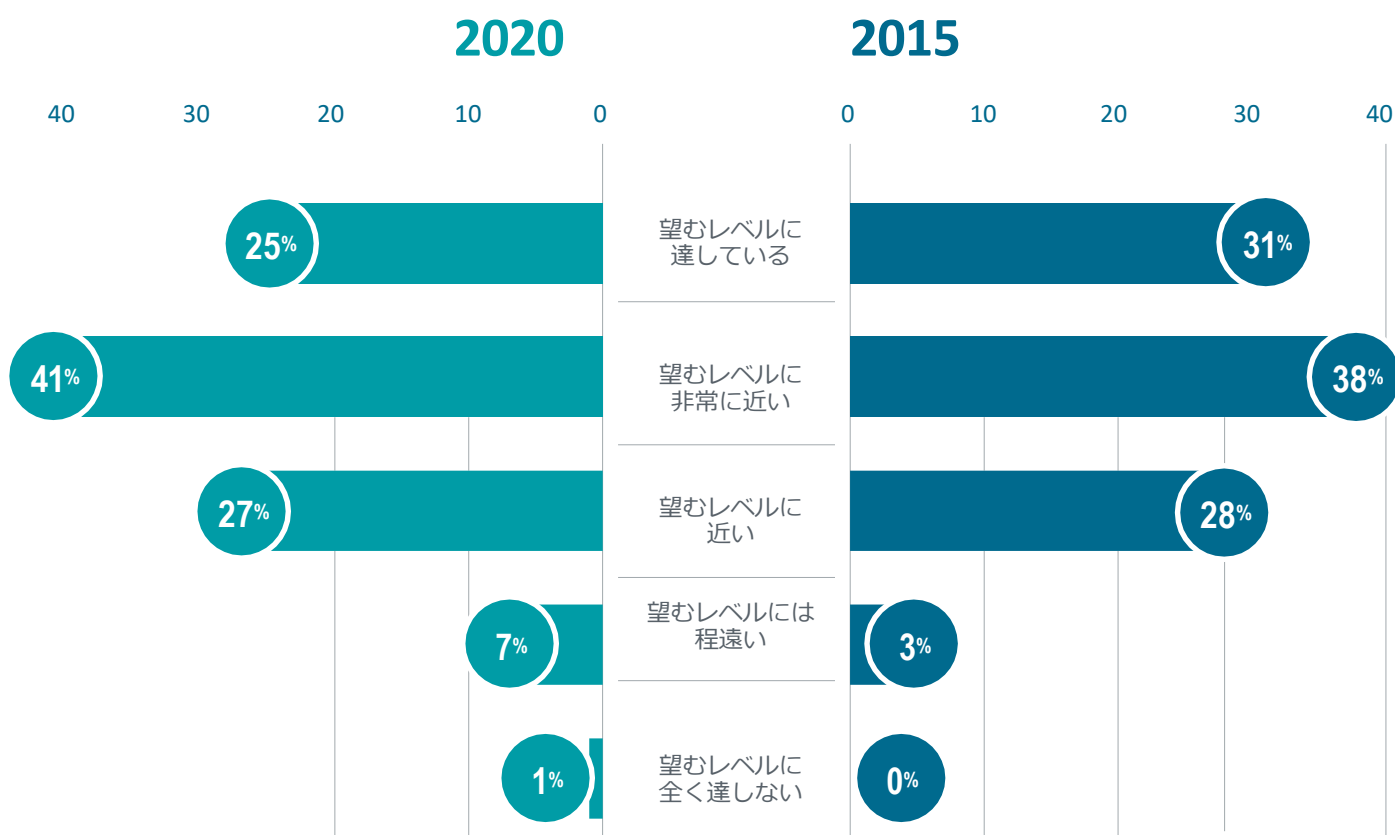


データ管理と分析に関して 「望むレベルに達している」とする 企業はわずか全体の4分の1

ほとんどのテクノロジートレンドと同様に、データを優先度の高いものとして認識するだけで、機能が向上するわけではありません。実際、企業はどれだけ遅れをとっているか認識し、後退してしまうことがよくあります。

CompTIAの最新のデータによると、自社のデータ管理と分析に関して「望むレベルに達している」とする企業は全体の4分の1に過ぎません。デジタルトランスフォーメーションのペースを考慮すると、ベストプラクティスが確立する以前に、この数字は低下してしまう可能性があります。

現在のデータ機能に関する見解



リアルタイムデータや大規模なデータセットから最大の価値を引き出すため、組織はデータスキルに焦点を置く必要があります。ますます多くの企業がデータを使い、業務を改善したり、顧客への理解を深めるなか、最新かつ改善されたスキルが成功へのカギとなります。そのようなスキルは、復元力のあるデータアーキテクチャの構築から、データ分析のスピードアップ、新たな洞察のためのデータマイニングまでビジネス上の幅広い問題に対処することができます。要は、新しいビジネスの通貨は、利害関係者のため価値を引き出すことができる新たな専門家を必要としているのです。

データが本当に重要なリソースであるならば、適切に扱う必要があります。データ機能に対する満足度が相対的に低いのであれば、改善の余地は十分にあります。データ機能をしっかり理解することで企業戦略を構築するための背景（コンテキスト）が見え、さらに、データに関連する主要な役割を理解することで、スキル構築のためのプロセスが生まれます。

多くの企業がデータを使って、業務を改善したり、顧客への理解を深めるなか、最新かつ改善されたスキルが成功へのカギとなります



データ機能に焦点を当てる： なぜ企業はデータを必要とするのか？

CompTIAのレポート「Functional IT Framework」では、エンタープライズテクノロジーの4つの柱、つまり企業が組織全体で独立した機能として管理しているIT領域について説明しています。

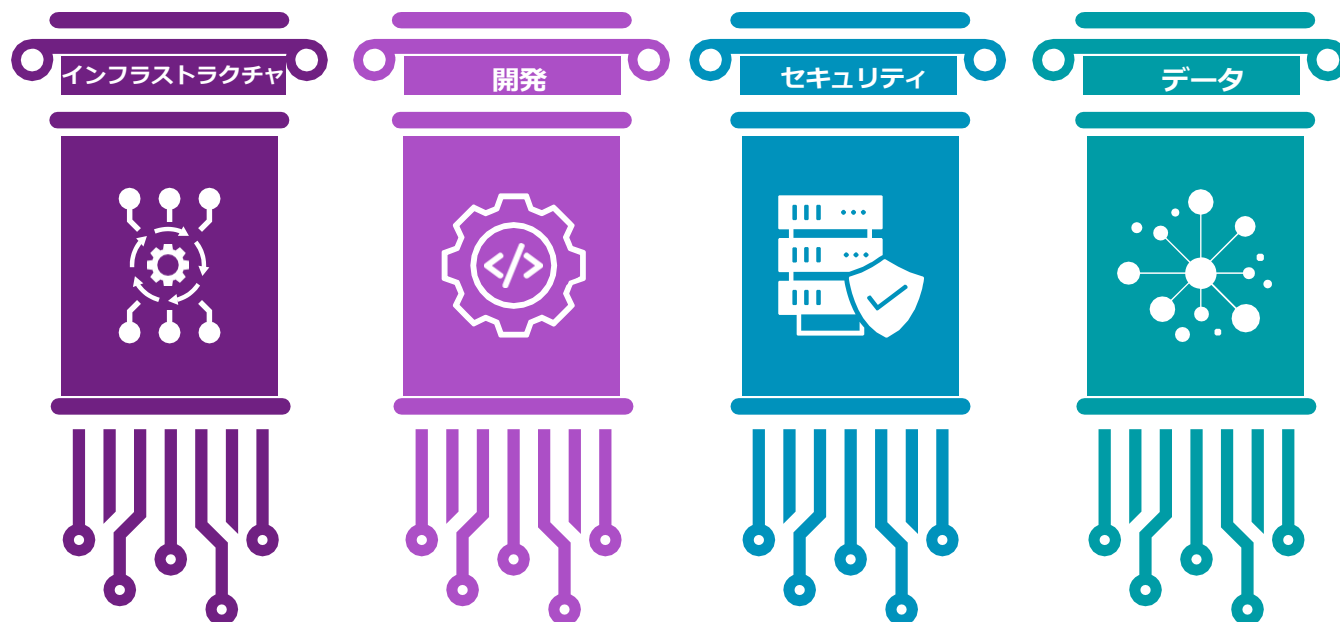
最も基本的な柱はインフラストラクチャで、システムをホストし、従業員をつなぐ物理的（またはバーチャル）ハードウェアコンポーネントです。この基盤の上に、機能を提供するアプリケーションを構築するのがソフトウェア開発です。これら二つの分野には、何十年にもわたって明確なスキル要件と運用プロセスがありました。

最近では、企業がデジタル資産の保護やリスク管理に重点を置くことから、サイバーセキュリティは専門化した分野となっています。歴史的に、サイバーセキュリティはインフラストラクチャ機能の中で扱われていたため、サイバーセキュリティが独立した機能となっても、これら二つの領域間には自然なつながりがあります。

データはこれら4つの柱の中で最も若いものですが、急速に成長しています。業界では、初めての最高データ責任者（CDO）が任命されたのは2002年だと言われており、NewVantage Partnersによる2018年の調査では、CDOを採用している企業数は2012年の12%から、62.5%に達したとしています。

データサイエンスは、テクノロジー分野で最も注目される話題の一つですが、ビッグデータ流行の最中に生まれて間もないものでもあります。企業がネットワークエンジニアやサーバー管理者のように、「データサイエンス」に匹敵する役割は少し前にはなかったのです。

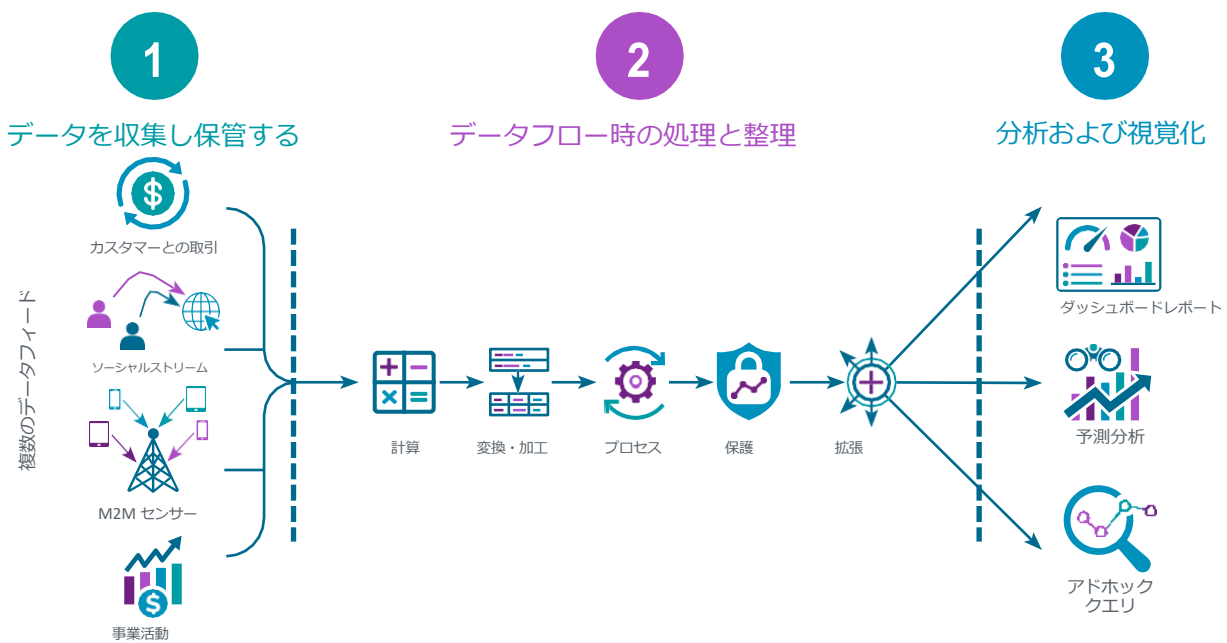
The Functional IT Framework



多くの企業では、データ機能はソフトウェア開発の派生物として始まります。開発に必要とされるスキルとクリティカルシンキングは、ビット (bits) とバイト (bytes) を処理する抽象的なコンポーネントを持つデータに変換されます。さらに、多くのデータスペシャリストは、PythonやJavaScriptなどのプログラミング言語を含むソフトウェア開発者と同じツールを使用しています。近年、企業が管理できるデータ量やデータの種類が増加していることから、特定の専門スキルに焦点が当てられていますが、より高度なアプリケーションに移行する前に構築しなければならない基盤があります。

近年、企業が管理できるデータ量やデータの種類の増加から、特定の専門スキルに注目が集まっています

データ使用における3つのステージ



その基盤は、組織内の全データの理解、データベースへの習熟、そして最も重要となる企業データの分析を含む、データの基本的な管理です。データ活動の全体を3つの基礎的な段階に分けることにより、組織は関連する疑問に到達します。当社のデータはどこから来るのか？様々なデータストリームをどのくらいの速さで処理すべきか？データを分析する際に、どのような洞察が提供されるべきか？これらの疑問を掘り下げることで、企業は、近代のデータ管理と分析に伴うテクニカル面での課題を理解することができます。

特に中小企業では、データベースに精通しているわけではなく、エクセルのスプレッドシートなどのシンプルなツールに頼って分析を行っているところが多いようです。しかし、企業のデータ高度化に伴って、データフローを効率的に管理し、分析するためのツールとスキルの両方が向上していきます。本格的なデータ活用の出発点となるのは、リレーショナルデータベースとSQL (structured query language 構造化クエリ言語) という二つの確立されたツールです。

リレーショナルデータベースは、データを有用な形式で保存するための従来のメカニズムです。データレコードは、各レコードの特定の部分の間に定義された関係をもって、データベースに置かれます。これらの関係を理解するためには、データベース内でデータが一貫して構造化されていなければなりません。そのため、保存できるデータの種類や実行できる分析に制限があります。リレーショナルデータベースのデータを操作するための最も一般的なツールは、SQLです。SQLには長い歴史があり、中央管理されたデータベースのスキーマや、インデックス付きのデータを操作することで開発者の間ではよく知られていますが、同時に限界もあります。

データ量の増加に伴い、モノリシックなリレーショナルデータベース上で動作するSQLアプリケーションのアーキテクチャは手に負えなくなります。さらに、収集されるデータの種類はもはや標準的なリレーショナルスキーマには適合しません。これらの問題に対処するために、二つのクラスのデータ管理ソリューションが登場しました。NewSQLとNoSQLです。

データツールセットの増加に伴い、組織全体でのデータスキルの需要も高まっています。ソフトウェア開発と同様に、データ機能もビジネスユニット内に設置されている場合があります。

NewSQLは、開発者がSQLインターフェースで培ったノウハウを活用しつつ、スケーラビリティとパフォーマンスの問題に直接対応しています。SQLシステムは、垂直方向のスケーリング（1台のマシンにコンピューティングリソースを追加すること）によって高速化することができますが、NewSQLシステムは、データベース自体のパフォーマンスを向上させたり、水平方向のスケーリング（新しいマシンを追加してリソースのプールを形成し、分散したコンテンツを可能にすること）を活用するために構築されています。NewSQLソリューションでは、SQL言語への接続を維持するとともに、企業はACIDに準拠したトランザクションを処理できるため、データベースの有効性が保証されます。

非構造化データを処理するために、多くの企業が従来のSQL製品とは異なるNoSQLソリューションを採用しています。NewSQLと同様に、NoSQLアプリケーションには、ドキュメント指向のデータベース、キーバリューストア、グラフデータベース、タブラーストアなど様々な種類があります。多くのNoSQLアプリケーションの基盤となっているのがHadoopです。Hadoopは、ビッグデータのプラットフォームとして機能するオープンソースのフレームワークで、ハードウェアリソースとエンドユーザーアプリケーションの橋渡しをする下位レベルのコンポーネントです。

データツールセットの増加に伴い、組織全体でのデータスキルの需要も高まっています。ソフトウェア開発と同様に、データ機能もビジネスユニット内に設置されている場合があります。部門ごとにデータに対するニーズは異なり、それぞれの社員が情報を精査し行動を決定することもあります。企業がテクノロジーガイドラインを維持しつつ、柔軟性を持たせるにはどうすればよいかを判断する際に、このようなデータ人材の分散は、データサイロを生み出し、他の組織的な課題につながります。

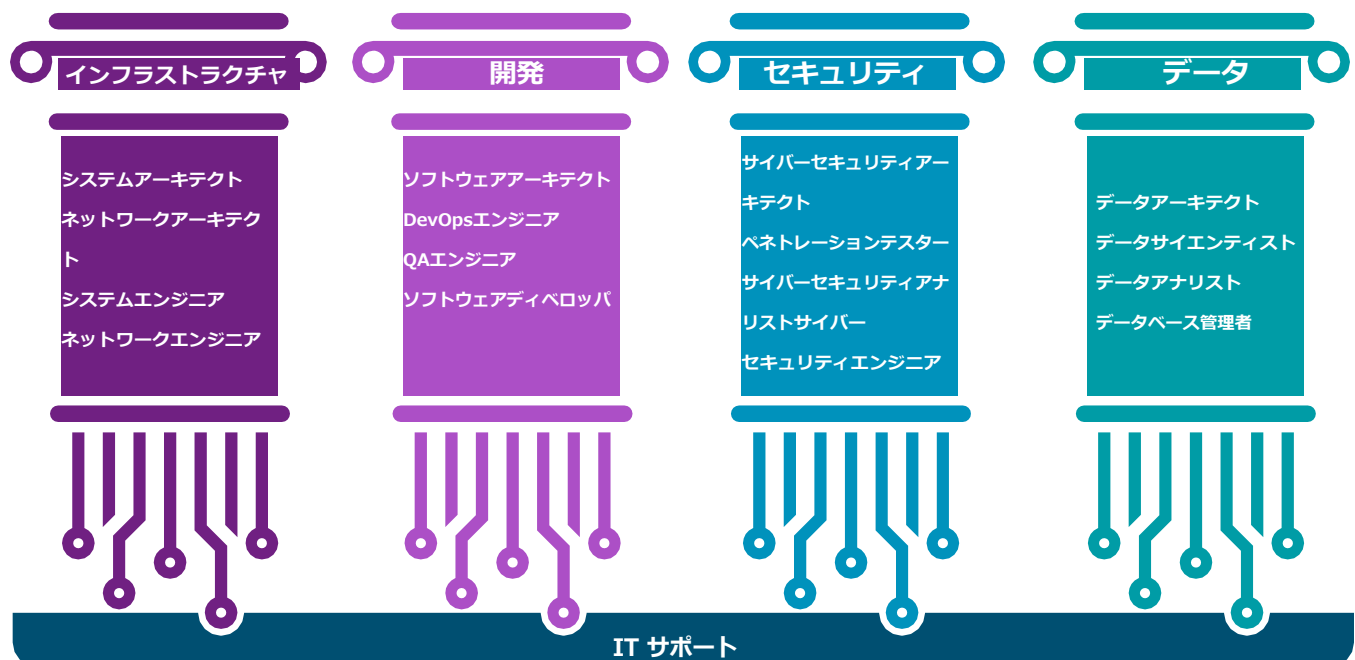
データチームの結成： 職務と必要なスキル

すべてのデータスキルがIT部門に集中している場合でも、複数のビジネスユニットに分散している場合でも、企業はデータチームの設立に向けて歩みを進めています。データチームの概念は比較的新しいものです。CISOは1990年代半ばに、CDOは2000年代初めに誕生したことを考えると、データチームがセキュリティチームに遅れをとっているのも納得できます。データ管理やデータ分析に特化した社員を社内に置いていると回答した企業は、全体の44%に過ぎません。すでに専任の社員を抱えている企業でも、新しいビジネス価値を生み出す分析に関するスキルを身につけたいというニーズは依然として高いことがわかっています。

データチームを構築する際、どのようなプレーヤーを追加する必要があるのかという初歩的な疑問があります。CompTIAは、Burning Glass Labor Insightsのデータを使用して、現在存在する17の異なるテクノロジーの役割を定義しました。この定義は、（役割の）数とスキルセットに基づいています。企業には、業種固有かつ既存の専門性に基づいて個々のニーズを持っているため、使用されている役割には無限のバリエーションがあります。

しかし、求人情報で使われている主要なスキルセットを調べると、現在の求人情報の大部分を占める、より小さな一連の職務に集約することができます。この17の職務は、固有の職務だけではなく、範囲や複雑さも伝えています。

基礎となるテクノロジーの職務



データ機能には、組織のデータチームを構成するための4つの役割があります。

- **データベース管理者:** データに関する最も伝統的な役割であるデータベース管理者（DBA）は、多くがソフトウェア開発部門からスタートしました（そして現在も多くのDBAがその部門に所属しています）。データベース管理者は、SQLやスクリプトを使ってリレーショナルデータベースに格納されているデータを操作することに重点を置いています。DBAはデータ構造をホストしているシステムを理解する必要があるため、この役割にはインフラストラクチャの要素も含まれることが多いです。一般的なスキルクラスターには、データベース管理、SQLデータベース、プログラミング、システム管理などがあります。
- **データアナリスト:** DBAは確立されているため、現在最も一般的な役割であるかもしれませんが、需要の面ではデータアナリストがリードしています。ほとんどの企業は、標準的なデータベース操作から得られる基本的な洞察を超えて、データとビジネスオペレーションとのつながりを求めています。「データアナリスト」と「ビジネスアナリスト」の境界は曖昧ですが、結局のところ企業はIT部門とビジネスユニットの両方でコアスキルを身につけることを求めています。一般的なスキルクラスターには、データ分析、ビジネスプロセスと分析、データビジュアライゼーションなどがあります。
- **データサイエンティスト:** 4つのテクノロジーの柱の中で、データ分野が最も新しい分野であるように、データサイエンティストは17の主要な職務の中で最も新しいものです。データアナリストとデータサイエンティストの主な違いは、統計的モデリングを用いて将来の予測を行うことです。最近では、マシンラーニングアルゴリズムが、データモデルの構築に重要な役割を果たすようになりました。この役割は、急成長している職種の一つとして挙げられることが多いのですが、それは真のデータサイエンティストの現在の基盤が非常に小さいことが大きな理由です。一般的なスキルクラスターには、統計学、ビッグデータ、マシンラーニングなどがあります。

企業が改善を求めているデータ機能の分野には、明確な優先順位があります。それは、アナリティクス機能の向上です。

- **データアーキテクト:** 4つの柱すべてにおいて、アーキテクトとされる役割の主な特徴は非常に似ています。アーキテクトは、様々な専門的な職務に就きますが、少なくとも各データ分野における実務経験を持ち、他のテクノロジー機能にもある程度精通しています。アーキテクトの役割は、適切なシステムを計画・構築することであるため、ビジネス上のトレードオフやリーダーシップについても精通しています。一般的なスキルクラスターには、プロジェクトマネジメント、システム設計や実装、コミュニケーションなどが挙げられます。

企業が改善を求めているデータ機能の分野を考えると、明確な優先順位があります。それは、アナリティクス機能の向上です。データ管理やデータベース管理（検索機能やリモートデータアクセス）に関連する改善も求められますが、要望の大半は分析能力の向上にあります。分析のスピードアップ、特定のビジネス活動に対する分析の適用、顧客行動の 패턴の発見などは、データアナリストが業務や意思決定の改善のために解決することができる課題の例です。

企業は多くの分野で改善を求めています

データ機能	現在うまくいっている	できているが改善したい	これからやってみいたい
入力されたデータのリアルタイム分析	40%	42%	13%
企業のデータソースを横断した検索機能	39%	37%	16%
Eメールマーケティングキャンペーンの効果	38%	37%	13%
ソーシャルメディアモニタリング	37%	33%	16%
ビジュアライゼーション機能	35%	37%	17%
ウェブサイトのトラフィックパターン	34%	38%	15%
顧客プロファイリングとセグメンテーション分析	33%	47%	11%
企業データへのリモートまたはモバイルアクセス	33%	43%	14%
予測分析機能	33%	42%	17%
指標と主要業績評価指標（KPI）の作成	33%	38%	17%
パターン認識	32%	38%	18%
リレーションシップ分析	29%	42%	16%

一般的に、データマイニングはデータ管理とデータ分析の2つの世界をつなぐ活動をいいます。データベース管理者は、環境の構築において主な責任を負いますが、アナリストもこれらの概念を理解する必要があります。特に、特に、あらゆる異種のデータセットがまとまりあるデータレイクに組み立てられていないビジネスの場合はなおさらです。アナリストは構造を理解した後、データを掘り下げてつながりやパターンを見つけることができます。

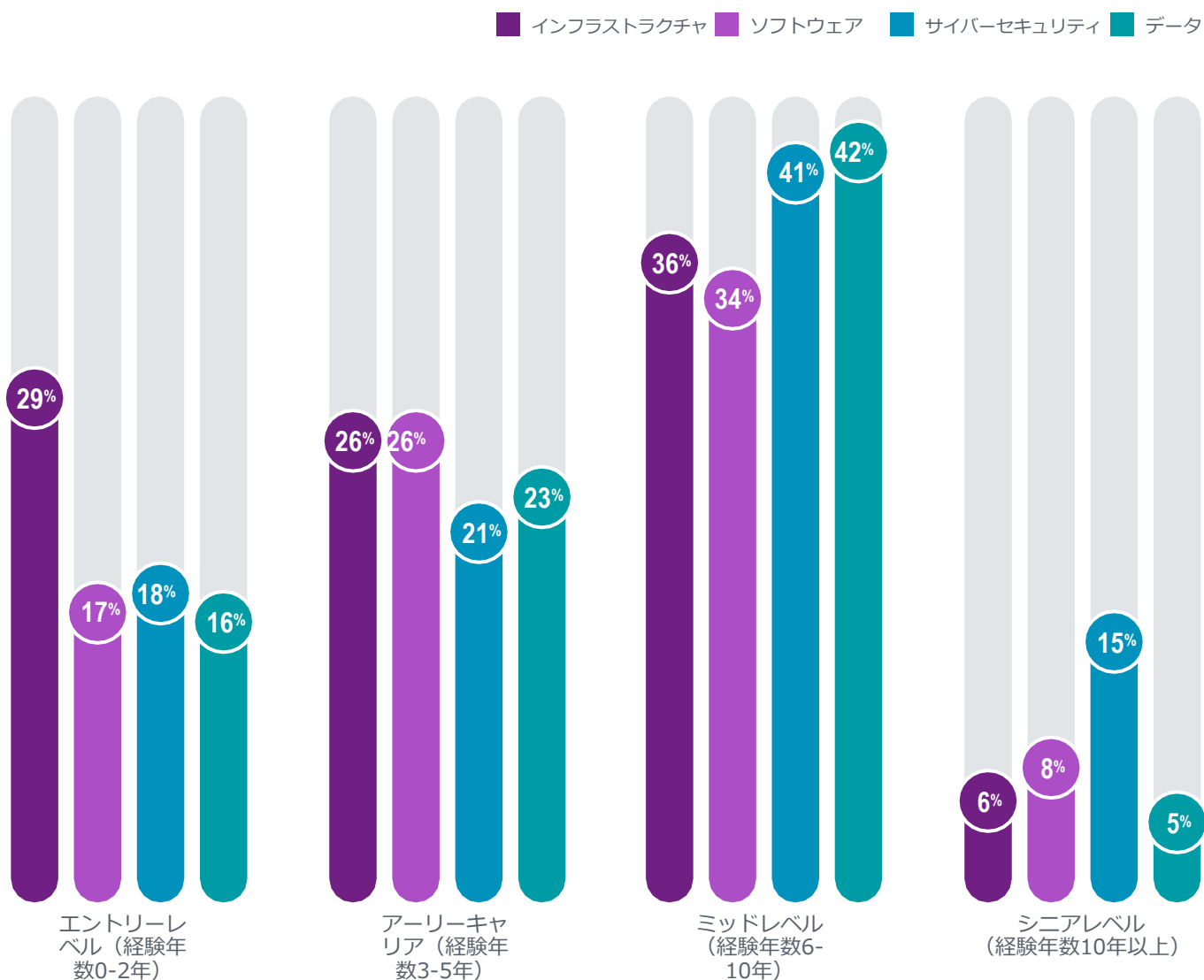
データビジュアライゼーションは、特に興味深い作業です。テクニカルな知識に加えて、ビジネスの知識やコミュニケーションスキルも必要になります。同様に、データガバナンスには、デジタルデータに関する規制や、州や国境を越えた規制の違いに関する専門知識が必要です。急速に変化するデジタル経済の中で、データ分析に必要なユニークスキルすべてを見つけることは、非常に大きな課題となっています。

データビジュアライゼーションは、特に興味深い作業です。テクニカルな知識に加えて、ビジネスの知識やコミュニケーションスキルも必要です。

データスペシャリストの発掘： データスキルの不足はありますか？

データ機能が比較的新しいことから、企業はエントリーレベルのポジションを強く求めているのではないかと考えられます。しかし、実際はそうではありません。事実、2021年に採用を予定している企業は、他のどの分野よりも中堅のデータスペシャリストを求めています。その理由の一つは、データやサイバーセキュリティという新興テクノロジー分野のチームは、既存のソフトウェアやインフラのチームから作られることが多いからです。しかし、専門的スキルに対する需要の高さは、簡単に解決できないパイプラインの問題を示しています。

ITプロフェッショナルを採用する際に求められる専門知識レベル



エントリーレベルの役割が明確でなければ、企業は、4年制の大学などのパイプラインから有能な人材を獲得し、彼らに業務経験やトレーニングの機会を与え、より高度なスキルを身につけさせるという、これまでのスキル開発の方法に頼ることができなくなります。高度なスキルに対する需要が爆発的に増えていることが、この問題をさらに悪化させています。

組織は、伝統的な4年制の学位プログラムを経験していない人材であっても検討する意欲を示しています。興味深いことに、4年制大学を卒業していないデータ人材を検討するという意欲は、ITサポート以外のどの分野よりも高く、エントリーレベルの職種としてのITサポートは、すでにブートキャンプなど様々なパスがある分野です。

4年制学位を持たない技術者採用の考察

	4年制学位のない人材を受け入れる可能性は低い/わからない	4年制学位のない人材を受け入れる可能性がある	4年制学位のない人材を受け入れる可能性が高い	可能性が高い + 可能性がある
ITサポート/ヘルプデスクポジション	17%	40%	43%	83%
データアナリスト/データベースポジション	21%	47%	33%	80%
ソフトウェア開発/ウェブ開発ポジション	25%	42%	33%	75%
サイバーセキュリティポジション	27%	40%	34%	73%
ネットワークエンジニア	28%	41%	31%	73%
CIO/ITマネージメント	31%	37%	33%	69%

企業が従来のパイプラインの考え方から脱却していく中で、どのような人材開発の取り組みであっても、短期的な人材の発掘と長期的な企業の成功の両方に貢献するものであることを確認する必要があります。特に、データ分野のように急速に進化するテクノロジー環境では、求職者、人事担当者、採用担当者など様々な人が関わるスキル開発プロセスにおいて、いくつかの重要な考慮事項があります。

まず、どのようなプロセスであっても、共通するテクニカル分野に関する「知識の証明」を含める必要があります。企業が自社の要件をより深く理解するためには、4年制大学の学位やIT認定資格など、専門知識を示す標準的な指標があることが望ましいでしょう。これらのクレデンシャルは多くの場合、基礎知識を確立するものであることから、候補者は独自の技術やベンダー固有の技術に関するトレーニングを必要とするかもしれません。しかし、長い目で見れば、強固な基礎知識を持った候補者を簡単に見つけることができるため、生産性の高い従業員を確保するまでの全体的な時間を短縮することができます。

二つ目は、スキル開発プロセスがこれまでの幅広い経験に適用されるべきという点です。データ分析のような分野では、理想的な人材はデータ構造や分析テクニック、運用方法や財務測定などのビジネススキルなど、様々なスキルを持っている可能性があります。トレーニングや認定資格は、前提条件となる情報はあっても特定分野の知識が不足している人材のスキルギャップを埋めることができます。

最後に、スキル開発プロセスは長期的な成功を目指したものでなければなりません。理想的なのは、より高いレベルで業務に慣れ、より早くキャリアアップできる人材を見つけることです。ペースの早いデジタル組織では、このような従業員は、自身の役割を受け入れ、成長の機会を見つけることで、ビジネスを前進させることができます。

データ分野では、時間が非常に重要となります。企業は膨大な量のデータを持っていますが、問題はそれをどう使うかです。データを迅速かつ効果的に分析する能力があれば、市場投入までの時間を短縮し、顧客満足度を高め、将来の成長につなげることができます。データを独立した機能とし扱い、データスペシャリストを育成することで、このプロセスを加速することができます。データ分析の迅速化を優先する組織は、スキルの発見とその開発の迅速化にも優先的に取り組む必要があります。

データ分野では、時間が非常に重要です。企業は膨大な量のデータを持っていますが、問題はそれをどう使うかです。



CompTIA.org

Copyright © 2021 CompTIA, Inc.. All Rights Reserved.

CompTIA is responsible for all content and analysis. Any questions regarding the report should be directed to CompTIA Research and Market Intelligence staff at research@comptia.org.