

《データセンター関連レポート Vol.2》 生成 AI がデータセンター市場にもたらす影響

2025 年 3 月 25 日

株式会社三井住友トラスト基礎研究所

新規事業開発室 研究員 宮本俊

- 近年話題となっている生成 AI は従来の AI をさらに発展させたものであり、データの分析や予測にとどまらず新しいコンテンツの生成が可能な技術である。
- 生成 AI は、多様なタスクに適応可能な基盤モデルが基になっており、必要に応じて下流のタスクに応用される点で、想定されている処理に特化して構築される従来の AI モデルと異なる。そのようなモデル構築のため、今までとはけた違いの量のデータを処理する必要があり、その基盤としてのデータセンターの需要が高まっている。
- そのような需要にこたえるためにより高性能なハードウェアの開発・導入が急速に進んでいるが、演算能力とともに電力消費の増加も著しく、必要電力の確保がデータセンター開発の最大の障壁となっている。
- さらに、消費電力の爆増により冷却システムの再考の必要性も出てきているが、導入コストの高さや既存の設備への新技術導入そのものが困難なことが新たなハードルとなっている。

昨今、生成 AI(Generative AI)に関連したニュースが連日報道され大きな注目を集めている。中国の AI スタートアップである DeepSeek が 2025 年 1 月に最新の AI モデル「DeepSeek-R1」を発表し話題となった。米国では Perplexity AI が「検索エンジンに代わる新たな情報収集ツール」として、AI による情報検索技術を発表した。また、日本国内でも政府が生成 AI の活用を推進する方針を発表し、企業や教育機関への導入が本格化してきている。さらに、OpenAI の ChatGPT、Google の Gemini といった生成 AI の開発競争が激化しており、独自のモデルが次々と発表されている。

生成 AI と従来の AI の大きな違いはそのアウトプットにあるとされる。従来の AI の主な役割は、与えられたデータに基づいた分析や予測を行うことであるが、生成 AI は与えられたデータを基に新しい文章、画像、音楽などのコンテンツを生み出すことが可能である。この特性を活かし、企業の業務効率化やクリエイティブな作業の支援、教育分野での学習補助など、多岐にわたる分野での応用が現実のものになってきている。

そのような革新的な技術の急速な発展と普及に伴って、高度な処理を行う基盤としてのデータセンターの需要がかつてないほど高まっており、市場も熱を帯びている。SoftBank が OpenAI とともに総額 5000 億ドルを投資し、AI システムの需要に応えるための共有コンピューティング施設の構築を進めると発表した「Stargate プロジェクト」が記憶に新しいが、その他にも英国では、政府が AI 投資プログラムを発表し、民間セクターから 140 億ポンドの資金提供を受けており、AI エネルギー評議会の設立、公共データの保護のための国家データライブラリの創設を推し進めるほか、公共セクターでの AI 活用にも取り組むとしている。

このように生成 AI の活用が本格化し、各国、各企業がその主導権をめぐるしのぎを削るなか、デジタルインフラの重要な基盤として拡大を続けるデータセンターの市場の成長はさらに加速するだろうと考えられている。その一方で、電力のひっ迫や環境への配慮等といった既存のデータセンターにおける問題点に加えて、より高度な冷却技術の必要性といった新たな課題もあり、どのようにして需要の拡大に対応していくのかがデ

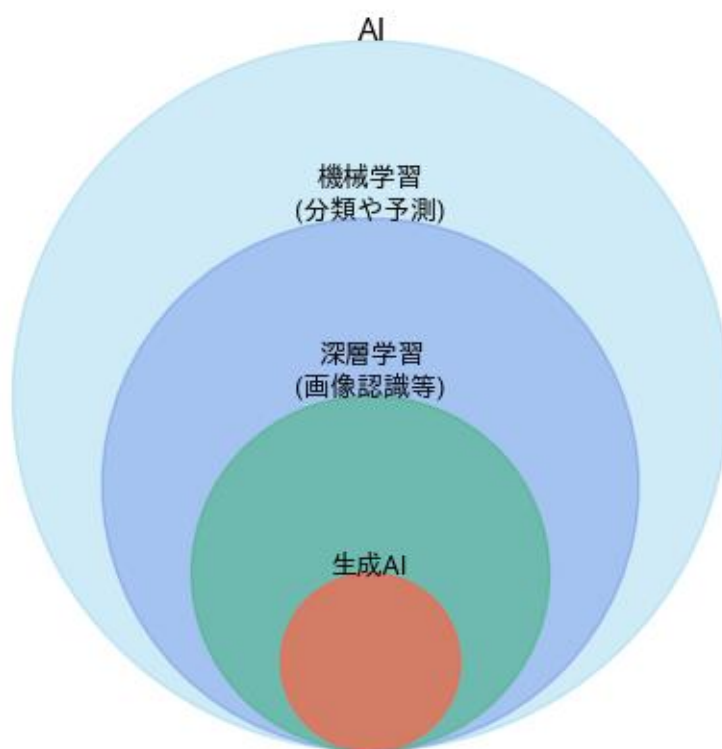
ータセンター開発のキーポイントになりつつある。

本稿では、まず生成 AI の基本原理や活用方法を整理したうえで、その発展によってデータセンター市場がどのように変化していくのかについて述べる。

生成 AI は多様なタスクに対応できる

先に述べた通り、生成 AI は新しいコンテンツを生み出すことができる点で、従来の AI と一線を画している。では、そもそも AI とはなんだろうか。一般的には、コンピュータに人間の知能・知性を模倣させようとする技術全体のことである。例えば、データの中からパターンを見つけ出し分類や予測を行う機械学習 (Machine Learning: ML) は AI 技術のうちの一つであり、画像認識等に用いられている深層学習 (Deep Learning: DL) は、ニューラルネットワークという機械学習の技術を応用していることからそれらのうちの一つとされている。そして生成 AI はその深層学習の技術の応用であり、予測や画像認識にとどまらず新たなコンテンツの創出を可能にしている。

図表 1 AI 技術の概念図

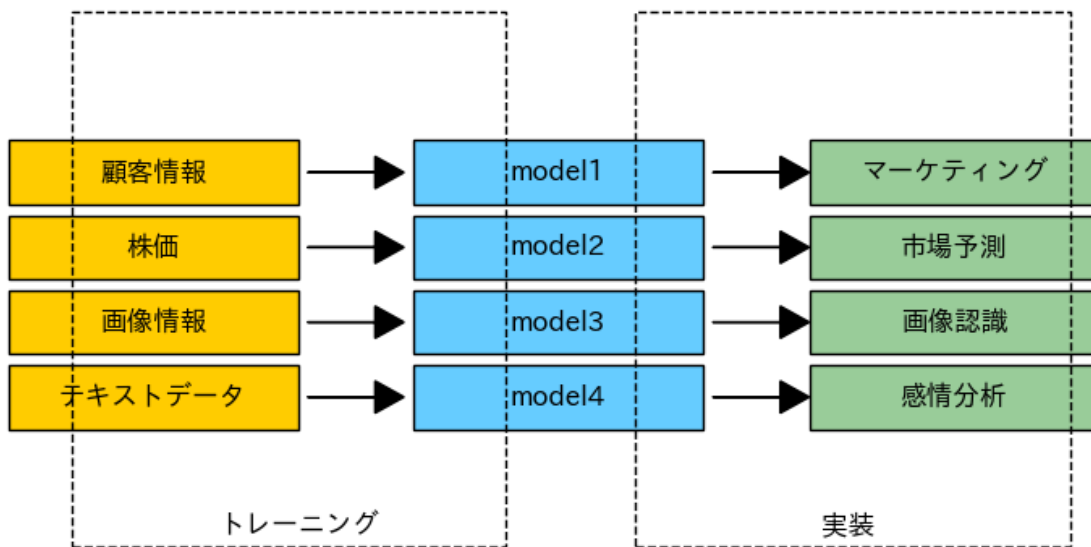


出所) 三井住友トラスト基礎研究所

生成 AI によるデータセンターへの影響を考察する時、その根本的な原因は生成 AI の基本的な原理に強く関連している。先に述べたように生成 AI は深層学習のうちの一つである。より具体的に言えば、生成 AI は大規模な深層学習モデルで大量かつ多様なデータを事前学習(Pre-Training)した基盤モデル(Foundation Model)がベースとなっている。この基盤モデルこそが生成 AI とこれまで考えられてきた AI の大きな違いとなっている。従来の AI モデルは次のように構築される。

- 処理したいタスクについてのデータセットを準備する。
- そのデータセットを用いて、個別のモデルをトレーニングする。
- トレーニング後、要求されたタスクに特化したモデルとして実装される。(例;メールフィルタ)

図表 2 従来のモデル構築の例



出所) 三井住友トラスト基礎研究所

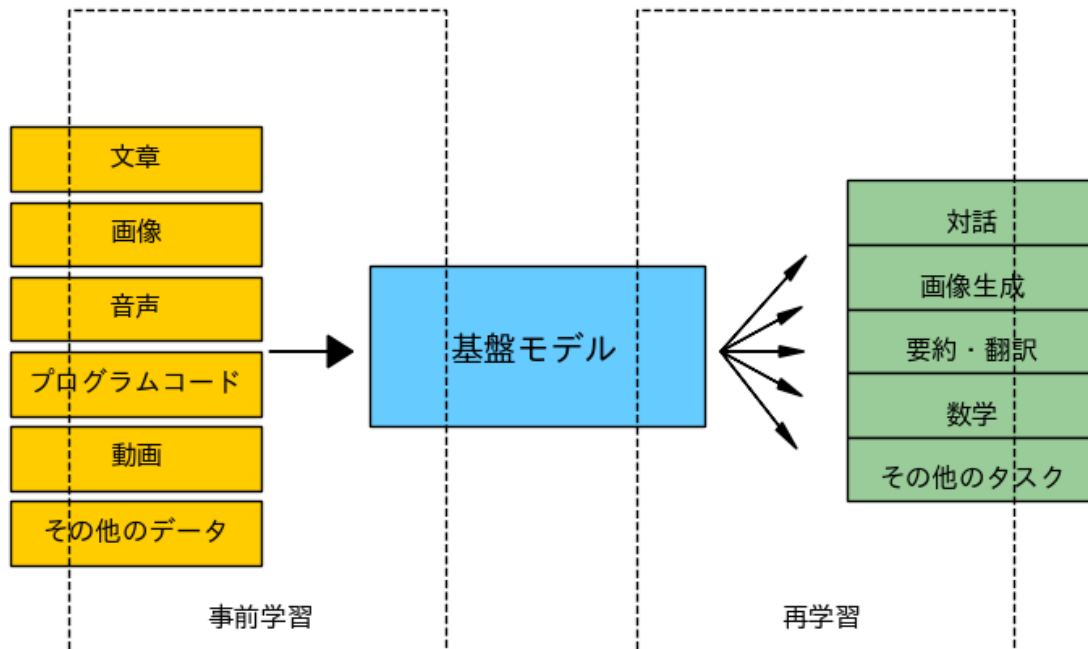
一方で、生成 AI につながる基盤モデルは異なる動作をする。

- 大量かつ文章や画像、音声などの多様なデータを準備する。
- このデータを用いて基盤モデルの事前学習を行う。
- 事前学習が完了した基盤モデルは、タスクごとに別々のモデルを訓練することなく多様なタスクに適応することができる。一つの基盤モデルで文章生成、要約、チャットボット等の役割を果たすことが可能である。
- すでに学習の済んだ基盤モデルに特定のデータセットを再学習(Fine Tuning)させることで、「法律文書の要約」や「プログラムコードの自動補完」といった、より明確で具体的な用途(下流タスク)に適応ができるようになる。

「事前学習後の基盤モデルと、下流タスク用にそれらを再学習させたもの」が生成AIと呼称されている。例えば、生成AIブームの火付け役となった ChatGPT の基盤モデルは、Web 上の大量のテキストデータ

を事前学習した GPT(Generative Pre-trained Transformer)であり文章生成や翻訳などの多様なタスクに対応ができる。ChatGPT はこのモデルを再学習によって会話のタスクに適応させたものである。

図表 3 生成AIのモデル構築の例



出所) 三井住友トラスト基礎研究所

汎用性と高い精度を担保するための莫大なデータ量

データセットの段階から特定のタスクに特化したモデルを考慮している従来の AI とは違い、生成 AI の基盤モデルは多様な目的に応用することを想定されている。そのような多目的・高精度なモデル構築を可能にしているのは、大量かつ多様なデータの事前学習である。

従来の AI モデルの実装例として「スパムメールの分類」を挙げると、モデルの構築にはデータセットとしておよそ 10 万のメール、データ量にしておよそ 500 メガバイトと考えられている。一方で、ChatGPT の基盤モデルである GPT3 はウェブ上のテキストデータや書籍、論文などから事前学習を行い、そのデータ量は 45 テラバイトに上るとも言われており、文字通り桁が全く違うものになっている。このような膨大なデータを処理する基盤としてもデータセンターの需要が急速に高まっていることが、供給動向の変化につながっている。

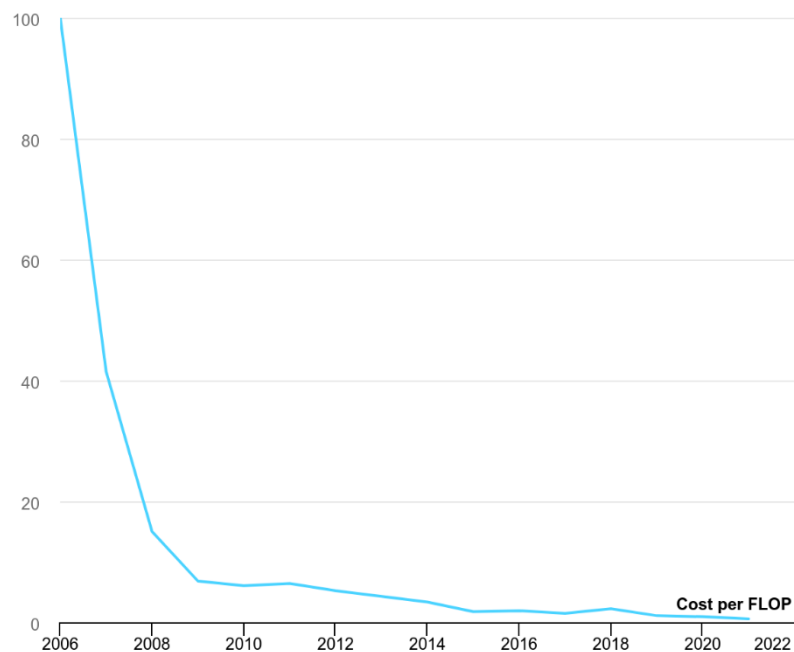
生成 AI の発展によってデータセンターの電力需給問題はより深刻に

データセンターは、我々の日常生活の中にデジタルツールが普及したこと等により、データのストレージ、分析、ネットワーク処理などの用途で現代社会にとって不可欠なインフラとなっており、生成 AI の登場以前から市場規模が拡大を続けている。その中で、近年データセンター市場では電力の供給という問題に直面している。より大きなデータ処理のためにハードウェアは高性能化し、その結果ラック(複数のサーバやストレージ、そ

の他ハードウェアを格納する設備)当たりの電力消費(以下、ラック密度)が上昇するケースが相次いでいる。その結果、海外では開発を認めない休止期間を設ける国・地域も出てきており、データセンター開発、運営の際の障壁の一つとなっている。

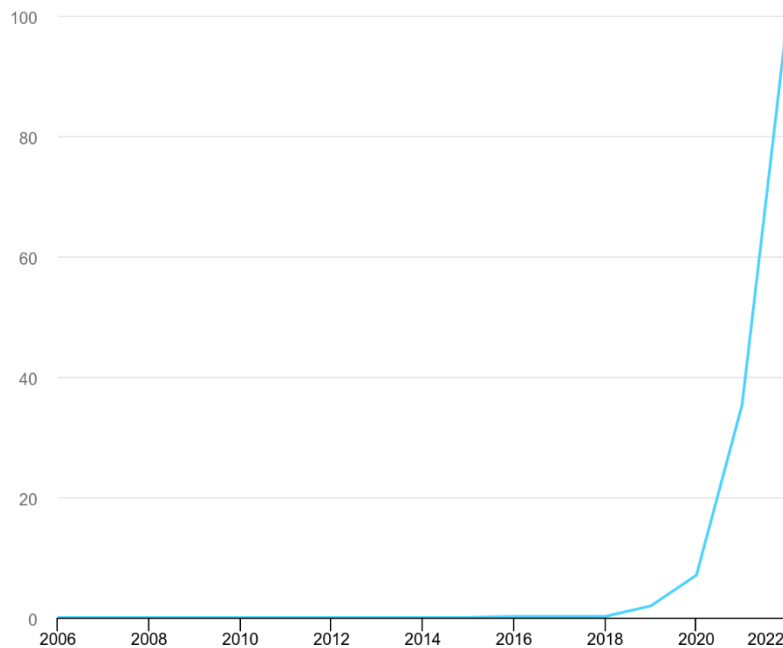
生成 AI の台頭によって、データセンターの役割は従来のデータのストレージおよびネットワーク処理から、より高度な AI 処理の基盤へと変化している。そのため、生成 AI の基盤モデルを構築するためには従来とは比較にならない量のデータを処理する必要があり、既存のデータセンターで用いられてきたハードウェアの処理能力を大きく超えてしまっているため、CPU (Central Processing Unit: 中央処理 (演算) 装置) と呼ばれる従来の演算機器ではなく、演算処理能力の高い GPU (Graphical Processing Unit: リアルタイム画像処理に長ける演算装置であり、並列処理に特化している) や機械学習や AI 用途に特化した TPU (Tensor Processing Unit) の導入が不可欠とされている。そのような需要を受けて NVIDIA や Google などによる AI 向け専用ハードウェアの開発競争が進んでおり、その処理能力は日進月歩で上昇していて、それに伴ってこれらを利用したモデルの構築もより高度なものになっている。2025 年 1 月に Meta が発表した、近年で供給が増えている 100MW のハイパースケール向けデータセンターをはるかに上回る 2GW (100MW の 20 倍) 以上のデータセンター開発計画では、AI 開発を目的としており、NVIDIA 製の AI 向け GPU を 130 万台導入すると報じられている。

図表 4 GPU 処理能力あたりの価格の推移 (2006 年を 100 とした場合)



出所) IEA (2024), Cost of graphics processing units, 2006–2022, IEA, Paris <https://www.iea.org/data-and-statistics/charts/cost-of-graphics-processing-units-2006-2022>, Licence: CC BY 4.0

図表 5 AI モデルのサイズと複雑さの推移(2022 年を 100 とした場合)



出所)IEA (2024), Artificial intelligence model size and complexity, 2006–2022, IEA, Paris

<https://www.iea.org/data-and-statistics/charts/artificial-intelligence-model-size-and-complexity-2006-2022>,

Licence: CC BY 4.0

こういったハードウェアの高性能化によって、今後はより一層、電力需給が新規供給の大きな障壁となるだろう。従来のデータセンターにおけるラック密度は概ね 5~20kW といわれており、一部の HPC (High-Performance Computing) と呼ばれる学術研究やシミュレーションなどに用いられる高性能ハードウェアで 40kW とされている。一方で、2024 年 3 月に NVIDIA が発表した新設計の GPU となる「Blackwell」シリーズでは、高負荷下での消費電力を示す TDP (Thermal design Power: 熱設計電力) は 1000W であり、対応するラックには合計で 72 基の GPU が搭載できるとされていることから、そのラック密度は 100kW 以上になると考えられる。もちろんこのスペックを必要とする施設がどれほどあるかは疑問が残るところではあるが、近い将来生成 AI の発展とともにこれに近いスペックのものの需要が増せば、電力不足がさらに強く懸念される。こういった事態に対し自社で発電設備を併設することで対策を講じている企業もある。株式会社ハイレゾは石川県の GPU 特化型データセンターにおいて、隣接地に自社の太陽光発電設備を設置し、再生可能エネルギーの比率を高める計画を進めている。また、Google や Meta などの大手テクノロジー企業も海外での新規供給時には再生可能エネルギーや原子力エネルギーの活用を視野に入れており、AI データセンターの過剰な電力需要に対応すると同時に、環境負荷低減と持続可能なエネルギー供給の実現を推し進めている。

高性能化によって生じる発熱問題

生成 AI がデータセンターにもたらすもう一つの大きな課題は、ハードウェアの冷却問題である。TDP は「コンピュータのパーツが発生させる熱の理論上の最大量であり、その冷却システムが放熱するようにデザインされている熱」のことを指すケースもある。すなわち TDP が大きくなるということは消費電力が増えることを指すのと同様に、発熱の量も大きくなることを意味するのである。そのため、生成 AI に適したハードウェアがより大きな TDP を持つことはデータセンターの冷却設備が従来のものでは対応しきれないことを意味している。冷却

能力の不足はハードウェアの動作不良やパフォーマンスの低下を招き、想定通りのパフォーマンスを発揮できなくなることから、生成 AI に適したデータセンターは、従来よりも高能力の冷却設備を導入する必要に迫られている。近年では、従来の空冷設備の代替となる技術として液体冷却に注目が集まっている。液体冷却技術は熱伝導性の高い冷却液を使用してハードウェアから発生する熱を効率的に吸収し外部に放出するという仕組みで、現在以下のような手法が考案されている。

技術	仕組み
液浸冷却 (Immersion Cooling)	<ul style="list-style-type: none"> ● ファンが不要で冷却エネルギー削減 ● シングルフェーズ液浸冷却: <ul style="list-style-type: none"> ➢ サーバを冷却液に浸し、ポンプで循環させる ➢ 液体は蒸発せず、循環しながら熱を拡散する ● デュアルフェーズ液浸冷却: <ul style="list-style-type: none"> ➢ 冷却液が発熱によって気化し、発生した蒸気を凝縮して液体に戻す仕組み
直接冷却 (Direct-to-Chip Liquid Cooling)	<ul style="list-style-type: none"> ● 各サーバのプロセッサに水冷プレートを装着し、冷却水を直接循環させる。 ● 温まった冷却水は熱交換器を通して冷却され、再度循環する。
リアドア式冷却 (Rear Door Heat Exchanger)	<ul style="list-style-type: none"> ● サーバラックの背面に設置された冷却パネルを通じて、ラック内の温風を直接冷却する。 ● 冷却水がパネル内を循環し、サーバが発する熱を効率的に除去する。 ● 空冷方式との併用も可能で、段階的な冷却設備の強化が可能。

出所) 三井住友トラスト基礎研究所

生成 AI の登場によりデータセンターは必要電力量の急増に加えて冷却設備の技術革新が求められ、新規供給のハードルが高まっている

今後、生成 AI 技術の進化によって、デジタルインフラとしてのデータセンターは高度な処理基盤としての役割も担うようになることから、その市場規模の拡大はさらに加速するだろう。それに伴い、電力需給の問題はより深刻さを増すだけでなく、新たな冷却技術導入の必要性も出てきている。しかし、冷却技術の革新は途上であり水冷の有用性は確認されているものの、導入コストが高く、従来のデータセンターに比べて初期投資が大きくなる。特に、液浸冷却は特殊な設備が必要であり、既存の施設への導入のハードルが高く、広範な普及にはより一層の技術革新が求められる。また、電力量の増大に伴い、環境問題やカーボンニュートラルなどへの対応にも迫られている。データセンターは、生成 AI の登場により、その役割だけでなく、冷却設備などの施設全体に求められる設備の技術革新の観点でも一つの転換点を迎えており、新規供給のハードルがより一層高まっている。

【お問い合わせ】新規事業開発室

<https://fofa.jp/smtri/a.p/122/>

1. この書類を含め、当社が提供する資料類は、情報の提供を唯一の目的としたものであり、不動産および金融商品を含む商品、サービスまたは権利の販売その他の取引の申込み、勧誘、あっ旋、媒介等を目的としたものではありません。銘柄等の選択、投資判断の最終決定、またはこの書類のご利用に際しては、お客さまご自身でご判断くださいますようお願いいたします。
2. この書類を含め、当社が提供する資料類は、信頼できると考えられる情報に基づいて作成していますが、当社はその正確性および完全性に関して責任を負うものではありません。また、本資料は作成時点または調査時点において入手可能な情報等に基づいて作成されたものであり、ここに示したすべての内容は、作成日における判断を示したものです。また、今後の見通し、予測、推計等は将来を保証するものではありません。本資料の内容は、予告なく変更される場合があります。当社は、本資料の論旨と一致しない他の資料を公表している、あるいは今後公表する場合があります。
3. この資料の権利は当社に帰属しております。当社の事前の了承なく、その目的や方法の如何を問わず、本資料の全部または一部を複製・転載・改変等してご使用されないようお願いいたします。
4. 当社は不動産鑑定業者ではなく、不動産等について鑑定評価書を作成、交付することはありません。当社は不動産投資顧問業者または金融商品取引業者として、投資対象商品の価値または価値の分析に基づく投資判断に関する助言業務を行います。当社は助言業務を遂行する過程で、不動産等について資産価値を算出する場合があります。しかし、この資産価値の算出は、当社の助言業務遂行上の必要に応じて行うものであり、ひとつの金額表示は行わず、複数、幅、分布等により表示いたします。