

FULLFACT独自分析レポート

AIを導入したのに売上も利益も増えない 原因と対策 2026

能力も価格も普及も改善。なのに約95%の組織が損益に効果を出せていない。乖離の原因と打ち手を一次情報で整理する

ABSTRACT

AIの能力は急上昇し、同じ性能を出すコストは約2年（約23か月）で280倍安くなり、利用も急速に広がっている。それでも組織の約95%は損益に測定可能な効果を出せていない。この乖離の原因は能力の不足ではなく、どの業務でどう測って効かせるかという実装の側にある。国内外の公開資料を再分析し、AIを業務で継続して使うための確認項目を整理する。

EXECUTIVE SUMMARY

主要論点

AIの能力は急上昇し、同じ性能を出すコストは約2年（約23か月）で280倍安くなり、利用も急速に広がっている。それでも組織の約95%は損益に測定可能な効果を出せていない。この乖離の原因は能力の不足ではなく、どの業務でどう測って効かせるかという実装の側にある。

01 能力は1年で跳ね上がった

2023年に導入された難関ベンチマークで、AIのスコアは1年でMMMU+18.8、GPQA+48.9、SWE-bench+67.3ポイント上昇した。実コードのバグ修正を測るSWE-benchでは、解けた課題の割合が2023年の4.4%から2024年の71.7%へ跳ね上がっている。

02 同じ性能が約2年で280倍安くなった

GPT-3.5並みの性能を出すための問い合わせコストは、2022年11月の100万トークンあたり20.00ドルから2024年10月には0.07ドルへ下落した。約2年（約23か月）で280倍超の低下である。ハードウェアのコストは年率30%で下がり、エネルギー効率は年率40%で改善している。

03 普及も止まらない

日本の個人の生成AI利用経験は2023年度の9.1%から2024年度の26.7%へ約3倍に伸びた。McKinseyの調査では、少なくとも一つの業務機能でAIを日常的に使う組織は88%に達している。能力も価格も普及も、すべて改善の方向に動いている。

04 それでも損益は動かない

MIT NANDAの調査では、企業が生成AIに300~400億ドルを投じたにもかかわらず、組織の約95%が測定可能な損益へのインパクトをまったく得られていない。能力と価格と普及が三拍子そろっても、財務成果はついてこない。これがパラドックスの核心である。

05 利益に効かせた企業はごく一部

McKinseyの調査では、全社の営業利益に何らかのインパクトがあったと答えた組織は39%にとどまり、そのうち大半はAIが寄与する利益は5%未満だとした。AIで全社利益に有意なインパクトを出せた高成果企業は約6%にすぎない。BCGの調査でも、AIから実際に価値を生み出している企業は26%で、残り74%は具体的な価値をまだ示せていない。

06 分かれ目は賢さではなく実装にある

価格が下がり性能が上がっても成果が出ないのは、能力の問題ではない。米国ではAI利用企業の57%がAIを3つ以下の業務機能に限定し、包括的に展開している企業は利用企業の4%にとどまる。どの業務で使い、何を測り、誰が確認し、どのデータを入れるかという実装の設計が、成果の有無を分けている。

READING NOTE

本ページの論点は、AIの利用率を増やすことそのものではなく、業務で継続して使える状態をどう作るかに置く。数字は導入済みか未導入かの判定ではなく、どの層で摩擦が起きているかを見るために使う。

IMPLEMENTATION VIEW

最初に確認する対象は、利用者数、契約ツール数、研修受講者数ではない。対象業務、利用データ、確認者、ログ、改善責任がそろっているかを見れば、PoCから運用へ進める余地が見える。

KEY DATA

約95% をどう読むか

生成AIに投資した組織のうち、損益への効果を測定できていない割合（MIT NANDA）を起点に、改善する指標と停滞する指標を分けて読む。

難関ベンチマークの1年の伸び	MMMU+18.8 / GPQA+48.9 / SWE-bench+67.3	2023年導入→2024年（ポイント）	Stanford HAI 2025 AI Index
SWE-bench 解決率	4.4%→71.7%	2023年→2024年（実コードのバグ修正）	Stanford HAI 2025 AI Index
同性能あたりの問い合わせコスト	\$20.00→\$0.07 / 100万トークン	2022年11月→2024年10月（約2年で280倍超）	Stanford HAI 2025 AI Index
推論価格の下落（全ベンチマーク中央値）	年50倍（2024年1月以降は年200倍）	同一性能あたり価格の年次下落率	Epoch AI
個人の生成AI利用経験（日本）	9.1%→26.7%	2023年度→2024年度（約3倍）	総務省
少なくとも1機能でAIを日常利用	88%	2025年調査（前年78%から上昇）	McKinsey
損益に測定可能な効果なし	約95%	生成AIに投資した組織（投資額300~400億ドル）	MIT NANDA
価値を生み出している企業	26%	残り74%は具体的価値が未達	BCG
米国でのAI利用の広がり方	57%が3機能以下 / 全社展開は4%	AI利用企業（米国）	U.S. Census Bureau

注: 各調査は対象国、母集団、調査時点、設問定義が異なる。数値は単純な順位表ではなく、AI実装のどこに摩擦があるかを読むための材料として扱う。

MEASUREMENT

個人利用、企業方針、職場利用、事業機能への展開は、それぞれ別の測定単位である。高い数字と低い数字を直接つなぐのではなく、どの層で設計が止まっているかを分けて読む。

INTERPRETATION

FULLFACTの分析では、AI実装の差はツール選定よりも、業務プロセス、データ、責任体制、学習機会、ガバナンスの有無として表れる。数字はその仮説を検証するための入口である。

ANALYSIS

利用率ではなく、業務に残る条件を見る

AIをめぐる数字は、相反する二つの方向に同時に動いている。能力は難関ベンチマークを1年で塗り替え、同じ性能を出すコストは約2年で280倍を超えて下がり、個人の利用経験は1年で約3倍に伸びた。改善を示す指標は、どれを見ても上を向いている。

ところが、企業の財務に表れる指標は動かない。生成AIに巨額が投じられても、損益に測定可能な効果を出せた組織はごくわずかとどまる。能力・価格・普及がそろって追い風になっているのに、成果だけが取り残されている。本レポートは、この乖離をAI生産性パラドックスと呼ぶ。

パラドックスという言葉を使うのは、原因が直感に反するからである。性能が足りないわけでも、価格が高すぎて手が出ないわけでも、誰も使っていないわけでもない。むしろ逆に、賢く、安く、広く使われている。それでも成果が出ないとすれば、原因はAIそのものよりも、AIを業務に組み込む側にある。本レポートでは、改善する指標と停滞する指標を国内外の一次情報から並べ、両者を分けるものが何かを整理する。

この乖離を読み解く鍵は、改善している指標と停滞している指標が、まったく別の場所を測っているという点にある。能力・価格・利用率はAIそのものの性質や入口の広がりを知る。一方で損益インパクトや高成果企業の割合は、AIを自社の業務に編み込んで初めて動く指標である。前者はベンダーの開発競争と個人の好奇心で勝手に伸びるが、後者は会社が業務を設計し直さない限り進まない。改善する指標だけを見れば「もう十分使える」と錯覚し、停滞する指標だけを見れば「まだ早い」と誤解する。

FULLFACTが注目するのは、米国データに表れた利用の浅さである。AIを使う企業の過半数が3機能以下に限定し、全社的に展開できている企業は利用企業のわずか4%。多くの企業にとってAIは、一部の人が一部の場面で試す道具のまま止まっている。性能が上がっても価格が下がっても、使う業務の範囲が狭ければ、全社の損益には現れない。95%が効果を出せないのは、AIの性能ではなく、使う業務の範囲の狭さに原因がある。

だから打ち手は「使う」から「どの業務でどう測って効かせるか」へ移る。日本のマクロ経済は、この問いを長く放置した先に何が起きるかを示している。個人の利用率が1年で約3倍に伸びても、それだけでは実質賃金や生産性の数字に転化しない。対象業務を選び、入れてよいデータを決め、確認の責任者を置き、成果を測る指標を先に定める。この実装の設計を四半期ごとに見直せる企業だけが、能力と価格の追い風を自社の損益に変えられる。

IMPLEMENTATION CHECK

実行に向けた確認項目

効
か
せ
る
業
務
を
選
ぶ

利用者を増やすのではなく、損益に直結する業務を一つ選んで深く使う。業務頻度、扱うデータ、確認者、成果指標が見える業務から始める。

測
る
指
標
を
先
に
決
め
る

導入の前に、何が何ポイント動けば成功かを定義する。利用者数や契約ツール数ではなく、業務の処理時間や売上・コストの変化で測る。

入
れ
て
よ
い
デ
ー
タ
を
線
引
き
す
る

どの情報をAIに渡してよいか、渡してはいけないかを業務ごとに決める。利用が広がるほど、この線引きの不在が成果と信頼の両方を損なう。

確
認
と
責
任
の
所
在
を
置
く

出力を誰が確認し、誤りを誰が直し、運用を誰が更新するかを業務ごとに決める。確認の空白が、PoCの出力品質と本番の成果を混同させる。

四
半
期
で
範
囲
を
広
げ
る

成果が出た業務の手順と失敗ログを残し、次の業務へ展開する。AI活用方針を年1回の文書ではなく、四半期ごとに見直す運用にする。

RISK & EDITORIAL VIEW

避けるべき進め方

AIツールの比較だけで導入判断を終えると、現場で使う条件が残らない。全社員向けの一般研修だけでも、業務ごとの確認責任や入力情報の扱いは決まらない。

PoCの出力品質を本番運用の成果と混同すると、確認者、ログ、更新者、KPIが空白になりやすい。AI実装では、使った結果をどう確認し、どう直し、どう続けるかを見る。

導入の主語を「全社」に置きすぎる進め方も避けたい。部門や業務によってデータの所在も確認責任も異なるため、規模や部門を問わず、業務単位で進め方を決めるほうが定着しやすい。

外部パートナーを使う場合も、ツールの納品で終わる支援は選ばない。業務の分解、確認ルール、運用責任まで残る形で進めると、社内に判断の型が蓄積する。

一次情報を並べると、能力も価格も普及も、数字はすべて改善しています。それでも約95%の組織が損益に効果を出せていないのは、AIの性能の問題ではありません。試した人の数ではなく、損益に触れる業務をどれだけ深く設計できたかで差がついています。利用率を追うのをやめ、売上やコストに直結する一つの業務を選んで深く使い、成果を測りながら四半期ごとに範囲を広げる。改善した能力と価格を自社の数字に橋渡しできる企業だけが、この追い風を成果に変えられます。

株式会社FULLFACT

代表取締役 足達彩人

SOURCES

主要出典

- Stanford HAI, The 2025 AI Index Report (Technical Performance / R&D)**
<https://hai.stanford.edu/ai-index/2025-ai-index-report/research-and-development>
- Epoch AI, LLM inference prices have fallen rapidly but unequally across tasks**
<https://epoch.ai/data-insights/llm-inference-price-trends>
- 総務省「令和7年版 情報通信白書 概要資料」**
<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r07/summary/summary01.pdf>
- McKinsey QuantumBlack, The state of AI in 2025: Agents, innovation, and transformation**
<https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>
- MIT NANDA, The GenAI Divide: State of AI in Business 2025**
https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf
- BCG, AI Adoption in 2024: Where's the Value in AI? (74% of Companies Struggle to Achieve and Scale Value)**
<https://www.bcg.com/press/24october2024-ai-adoption-in-2024-74-of-companies-struggle-to-achieve-and-scale-value>
- U.S. Census Bureau, The Microstructure of AI Diffusion**
<https://www.census.gov/library/working-papers/2026/adrm/CES-WP-26-25.html>

注: 出典ごとに対象国、母集団、調査時点、設問定義が異なる。本文では単純な順位表ではなく、AI実装の論点を読む材料として扱う。

NEXT STEP

AI実装の論点を、対象業務・データ・責任体制まで整理する。

FULLFACTでは、本レポートで提示した観点をもとに、AI実装で最初に扱う業務、利用できるデータ、責任体制、確認ルールを整理する無料顧問制度を10枠限定で案内しています。

必要に応じて、業務・データ・組織体制を確認するAI実装診断も活用できます。

<https://fullfact.net/contact>

PRIMARY OFFER

10枠限定 無料顧問制度

1. 対象業務の選定
2. 利用データの棚卸し
3. 責任者・確認ルールの整理

SECONDARY

AI実装診断