レガシーの正体は、古いコードではなかった

システム開発・レガシー刷新のAI徹底活用

コードは読める。理由は、残っていない。
そこに、刷新が進まない本当の理由がある。

株式会社TMSI Partners ／ 2026年7月

この資料は、銀行編の5業務のうち「システム開発・レガシー刷新」を扱う最終編です。

当社Insightsの記事「金融機関がAIを『徹底活用』すると、業務はどこまで変わるのか。」では、
銀行の5業務・証券の4業務について到達点を整理しました。
本資料はそのうち銀行5「システム開発・レガシー刷新」を取り出したものです。銀行1〜4が「AIで各業務がどう変わるか」を扱ったのに対し、本編が扱うのはその変化を実装する能力そのもの——つまり銀行1〜4の前提条件にあたります。だから銀行編の最後に置きました。

本資料が扱うこと

- レガシーが「いつかやること」から「間に合うかどうか」に変わった理由
- コーディング支援AIが実際に速くした工程と、速くしなかった工程
- AI生成コードの安全性が2年間動いていないこと、そして言語別の偏り
- レガシーの定義を「古いコード」から「決定の履歴が失われたシステム」へ書き換える
- 自行の成熟度を測る4段階と、最初の90日でやること・やらないこと

本資料が扱わないこと

- 特定ベンダー・製品・開発ツールの比較評価および推奨
- 移行先アーキテクチャの選定指針（現行仕様の把握が先行するため）
- リテール・顧客接点、融資・審査、事務、リスク・AMLの各業務（別稿）
- 一次資料で裏付けの取れない数値（p.24に不採用理由を開示しています）

A	発表主体の一次資料で確認できたもの。本資料で図表にしている数値はすべてこの等級です。
B	報道ベースでのみ確認できたもの。本文中で出所を明示した箇所に限り言及し、図表には用いません。
C	原典を特定できなかったもの。掲載していません（p.24に一覧）。

A級数値からの単純計算は「当社試算」と明記し、図中では sand_deep の網掛けで区別しています。

本資料は経済産業省・情報処理推進機構（IPA）・金融庁・JUAS等の公表資料および各社の公式発表、ならびに公開された研究成果をもとに、株式会社TMSI Partnersが独自に整理・分析したものです（2026年7月時点）。

なぜ、コードを書くのが速くなっても、刷新は進まないのか。

2025年から2026年にかけて、コーディング支援AIは実務に定着した。DORAの調査では回答者の90%がAIを業務で使い、80%超が生産性の向上を実感している。金融庁のAIディスカッションペーパーも、システム開発分野を準備段階から導入済みの段階へ進んだと整理し、この分野で生成AIの活用が大きな成果をあげているとの指摘が聞かれる、と記した。

にもかかわらず、レガシー刷新は進んでいない。経済産業省とIPAの総括レポートによれば、61%の企業がいまだレガシーシステムを保有し、大企業に限れば74%。メインフレームからの移行を進める企業のうち、過半数が移行先すら決まっていない。

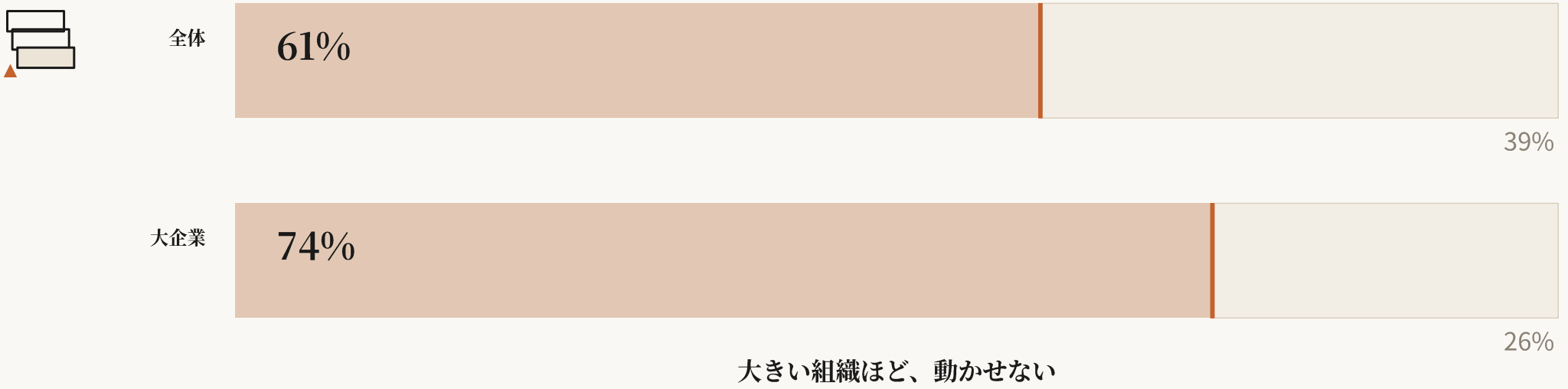
普通に考えれば逆である。開発が速くなれば、刷新も速くなるはずだ。ならない。そのズレの正体を解くのが、この資料である。

速くなったのは、書くことだけだった。
銀行の刷新を止めているのは、読むことと、確かめることだ。

そしてもっとも深いところには、コードを読んでも復元できないものが埋まっている。

まだ、6割が古い仕組みの上にいる。しかも大きい組織ほど動かせない。

経済産業省とIPAのレガシーシステムモダン化委員会は、2025年5月の総括レポートで、レガシーシステムを保有する企業が61%、大企業に限れば74%にのぼると報告した。メインフレームからの移行を進める企業のうち非メインフレーム環境へ向かうのは52%だが、そもそも過半数が移行先を決められていない。



これは一部の遅れた会社の話ではない。規模が大きく、止められない業務を抱えている組織ほど、この位置にいる。

経済産業省・情報処理推進機構（IPA）「DXの現在地とレガシーシステム脱却に向けて」レガシーシステムモダン化委員会総括レポート（2025年5月28日）。

予算の四分の三は、現状維持に消えている。

JUASの企業IT動向調査2025によれば、2024年度のIT予算のうち運用保守（ランザビジネス）が76.2%を占め、新規投資（バリューアップ）は23.8%にとどまる。有効回答981社の実績値である。当社試算では、新しいことに使える額は現状維持に使う額の約0.31倍にすぎない。

IT予算の内訳（2024年度実績）



新規投資

23.8%

運用保守

76.2%

TMSI試算

0.31倍

新しいことに使える額は、現状維持に使う額の約3分の1

23.8 ÷ 76.2 = 0.312

刷新の原資は、刷新しないことに食われている

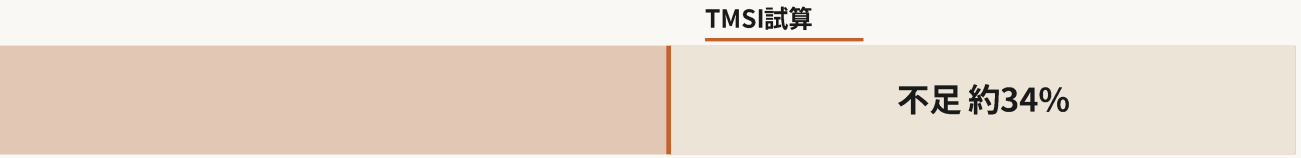
予算が足りないから刷新できないのではない。刷新していない古い仕組みを支えることに、予算が消えている。

一般社団法人日本情報システム・ユーザー協会（JUAS）「企業IT動向調査2025」（2025年4月／調査期間2024年9月6日〜10月28日、配布4,500社・有効回答981社・回収率21.8%）。0.31倍は当社試算（23.8 ÷ 76.2）。

足りないのは、手ではなく設計者だ。

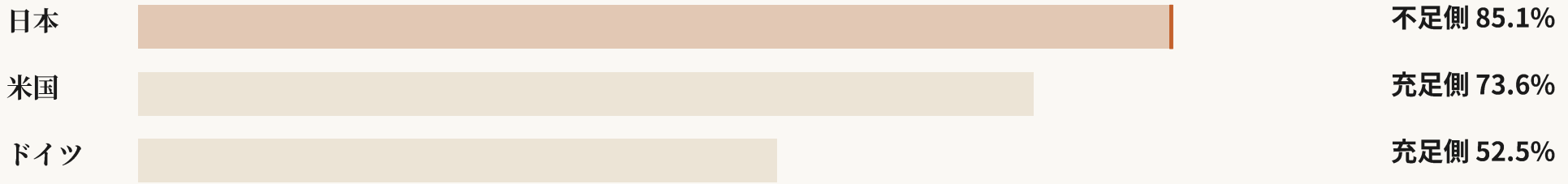
IT人材の供給は需要の66%。当社試算では約34%が不足している計算になる。ただし本質は量ではない。同レポートが特に不足していると指摘するのは、ビジネスアーキテクト、ITアーキテクト、データサイエンティストといった上流の人材である。設計を決める人がいない状態で、実装だけを速くしても行き先が決まらない。

IT人材の需給



不足しているのはビジネスアーキテクト、ITアーキテクト、データサイエンティスト

DX人材の充足感（集計区分が異なる）



日本＝不足している回答比率、米国・ドイツ＝過不足なし＋過剰。数値の直接比較はできません。

そしてこの不足感は、国際的に見ても日本が突出している。

経済産業省・情報処理推進機構（IPA）「DXの現在地とレガシーシステム脱却に向けて」（2025年5月28日）。国際比較は情報処理推進機構（IPA）「DX動向2025」（2025年6月26日／調査期間2025年2月上旬〜3月下旬、日本・米国・ドイツ）。不足率34%は当社試算（100 − 66）。日本と米国・ドイツは集計区分が異なるため、数値の直接比較はできません。

期限を決めているのは、
もう自社ではない。

これまでのレガシー議論は「いつかやらなければならない」だった。いま起きている変化は、やる時期を自分で決められなくなったことである。

金融庁は2025年6月のITレジリエンス分析レポートで、こう書いた。

> **メインフレームの事業から撤退することを公表する中、
> 金融機関が勘定系システムの更改を迫られており**



「迫られており」。監督当局が、金融機関の刷新時期を決めているのがハードウェアとソフトウェアのライフサイクルであると、自ら記述している。刷新の意思決定が、経営の計画からベンダーの製品計画へ移った。

国内では、大手ベンダーがメインフレーム事業から段階的に撤退する方針が報じられている（販売終了2030年度・保守終了2035年度と報道されているが、本資料では報道ベースの情報として扱い、年限を数値として掲載しない）。重要なのは特定の年ではなく、期限が外から来るという構造のほうである。

計画を立てる余裕があるうちに動くか、期限に追われて動くか。違いはそこだけになった。

金融庁「金融分野におけるITレジリエンスに関する分析レポート」（2025年6月30日公表）p.21。メインフレーム事業の終了時期に関する記述は報道ベースであり、ベンダーの公式発表として確認できていないため、本資料では年限を数値として掲載していません。

待つことにも、コストがある。

2024年度に報告されたシステム障害

約1,800件

TMSI試算

1日あたり 約4.9件

原因の約7割



1,800 ÷ 365 ＝ 4.93

金融庁のITレジリエンス分析レポートによれば、2024年度に報告されたシステム障害は約1,800件。当社試算では1日あたり約4.9件にあたる。

注目すべきは原因の内訳である。ソフトウェア障害と、管理面・人的要因。この二つで全体の約7割を占める。ハードウェアの故障が主因ではない。老朽化した設備そのものより、それを人手で支え続けている運用の側で起きている。

刷新を先送りするという判断は、無コストではない。支える負荷は毎年積み上がり、その一部は障害という形で表に出る。そして障害対応に投入される人員は、刷新に投入できたはずの人員と同じ母集団である。

> **先送りは、選択肢のひとつではない。**
> **毎年、代金を払っている。**

ここまでが第一幕の事実である。次に、その解決策として期待されているものを見る。

金融庁「金融分野におけるITレジリエンスに関する分析レポート」（2025年6月30日公表）脚注5および図表1。1日あたり約4.9件は当社試算（1,800 ÷ 365）。

やるかどうかではなく、
間に合うかどうかの問題になった。

ここまでの五つの事実を並べる。6割がレガシーを保有し（大企業では7割超）、予算の四分の三は現状維持に消え、上流の人材は需要の3分の1が欠けている。そこへ、ベンダーの製品計画という外部の期限が来た。そして先送りの代金は、年間約1,800件の障害という形で支払われている。

この状況で、誰もが同じことを考える。「開発をAIで速くすれば、間に合うのではないか」。

妥当な発想である。そして実際、コーディング支援AIはすでに広く使われている。だが、その期待が銀行のレガシー刷新にそのまま当てはまるかどうかは、別の話だ。

では、AIで開発を速くすれば、刷新は間に合うのか。

第二幕では、この問いに、公開されている実測データだけで答える。

90%が使い、80%が速くなったと感じ、30%が信じていない。

DORAが2025年9月に公表した「State of AI-assisted Software Development」は、約5,000名の回答と100時間を超えるインタビューに基づく、この分野で最大規模の調査のひとつである。そこには噛み合わない三つの数字が並んでいる。

90%

AIを業務で使っている

導入の議論は終わった。すでに現場の標準になっている

80%超

生産性が上がったと感じている

体感としての効果は、広く共有されている

30%

生成されたコードを信頼していない

「ほとんど、あるいは全く信頼していない」と回答

使っている。速くなったと感じている。それでも信じていない。

使っている。速くなったと感じている。それでも、生成されたものを信頼していない。この三つが同時に成立している状態を、どう読むか。

DORA「State of AI-assisted Software Development 2025」（Google Cloud、2025年9月23日／回答者約5,000名、100時間超のインタビューを含む）。

AIは増幅器である。

同じ調査は、AIの導入がスループットとプロダクト成果には正の関係を持つ一方で、デリバリーの安定性には負の関係を持つことを報告している。速くなった分だけ、壊れやすくもなっている。

DORAは、この現象に名前を付けた。

> **AIは増幅器（amplifier）であり、**
> **組織の既存の強みと弱みを拡大する。**



この一行が、本資料でもっとも重要な指摘である。強い組織はもっと強くなり、弱い組織はもっと速く壊れる。AIは差を埋める道具ではなく、差を開く道具だということだ。

したがって「AIを入れれば追いつける」という前提は成り立たない。テストが薄いチームは、薄いまま大量のコードを抱えることになる。設計の意図が共有されていない組織は、共有されないまま実装だけが増える。増幅されるのは、いま持っているものである。

ならば問うべきは「AIを入れるか」ではない。増幅されて困るものが、自組織に何があるか、である。

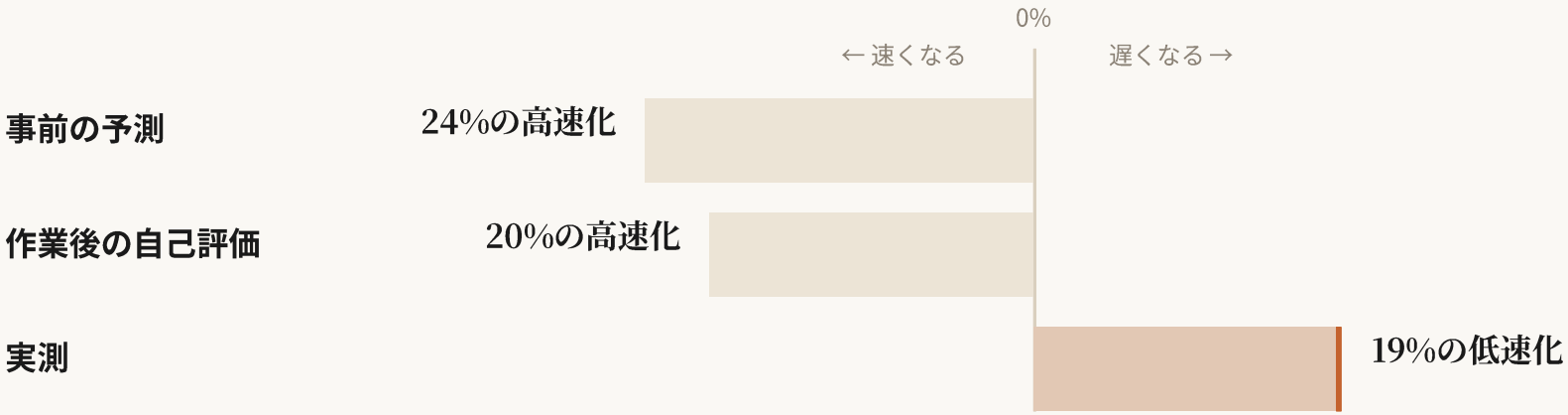
DORA「State of AI-assisted Software Development 2025」（Google Cloud、2025年9月23日）。増幅器（amplifier）は同レポートの表現。

熟練者ほど、速くなったと感じて、遅くなっていた。

METRが2025年7月に公表したランダム化比較試験は、平均22,000スター超・100万行超のリポジトリに複数年貢献している熟練OSS開発者16名に、246の実課題を割り当てたものである。結果は、AIを使える条件で完了時間が19%増加した。開発者は事前に24%速くなると予測し、作業後の自己評価でも20%速くなったと答えていた。

熟練OSS開発者16名・246課題のランダム化比較試験（2025年前半のAIを使用）

2025年7月公表



速くなったという実感と、実際に速くなったかどうかは、別の変数である。

TMSI試算

予測と実測の差 ＝ 約43ポイント

当社試算では、事前の予測と実測の差は約43ポイント。
この乖離のほうが、19%という数字そのものより重要である。

METR「Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity」（2025年7月10日／arXiv:2507.09089）。約43ポイントは当社試算（24 −（−19））。本研究には2026年2月に公表された続報があります（p.13）。

ただし、この研究には続きがある。

2026年2月、METRは第2次研究の設計を変更したと公表した。2025年8月に開始した追試で、深刻な選択バイアスが生じたためである。参加した開発者の30〜50%が、AIなしではやりたくない課題を、そもそも実験に提出しなかった。

そして実験に戻ってきた開発者に限った結果は、マイナス18%——つまり18%の高速化だった（信頼区間はマイナス38%からプラス9%と幅が広く、確定的な結論ではない）。

METR自身は、こう書いている。「2026年初頭の現在、開発者は2025年初頭の推定よりも高速化されている可能性が高い」。

研究者本人が更新した見立てを、同じ重みで載せる。

都合のよい数字だけを選ばないこと。それが、この種の資料が信頼される唯一の条件である。

METR「We are Changing our Developer Productivity Experiment Design」（2026年2月24日）。信頼区間の広さから、第2次研究の数値は方向性を示すものであり、確定値として扱うべきではありません。

問いを、立て直す。

AIは開発を速くするのか、遅くするのか。この問い方では、答えが出ない。研究ごとに結果が割れ、同じ研究でも年をまたぐと結論が動く。それは測定が下手だからではなく、問いが粗すぎるからである。

開発という営みは、単一の工程ではない。少なくとも三つに分かれる。書く。読む。確かめる。この三つは、必要な能力も、AIの効き方も、まったく違う。

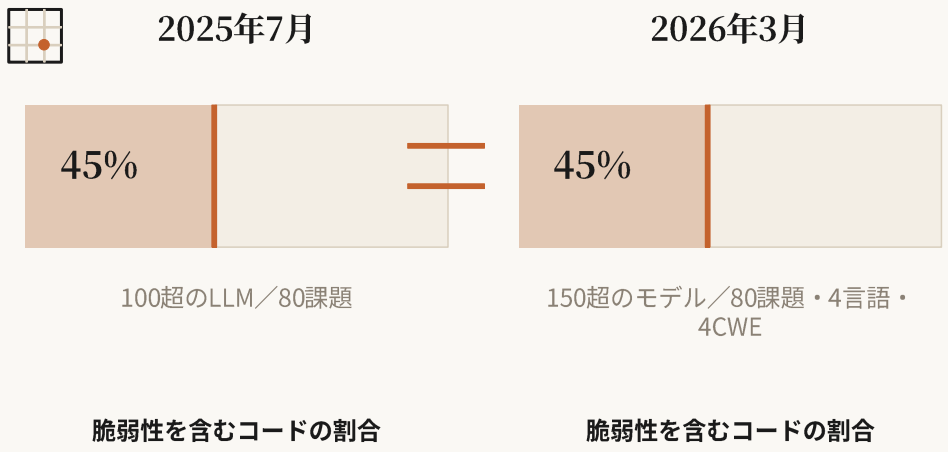
そして銀行のレガシー刷新において、詰まっているのがどの層かを特定しないまま投資すると、投資した層だけが速くなり、全体は変わらないという結果になる。DORAが観測した「スループットは上がるが安定性は下がる」は、まさにその状態の記述である。



問うべきは「速くなるか」ではない。「どの工程が速くなったか」である。

その前に、もうひとつ確認しておくべき事実がある。速くなった第一の層——書くこと——の、品質の話だ。

2年間、45%のまま。



Veracodeは2025年7月、100を超えるLLMに80の課題を解かせ、AI生成コードの45%に脆弱性が含まれることを報告した。そして2026年3月の更新版では、150を超えるモデルで再測定して、45%のまま——2年間まったく改善していないと結論している。

同レポートは、より大きなモデルが有意に優れているわけではない、とも書いた。モデルは賢くなった。安全にはならなかった。

この事実の意味は、待っていても解決しない、ということである。次世代モデルを待つという判断は、少なくともセキュリティの観点では合理的な待ち方ではない。解決策があるとすれば、モデルの側ではなく、
生成されたものを検証する側にある。

**モデルは賢くなった。
安全にはならなかった。**

**モデルの進歩では解決しない。
ならば、受け止める側を作るしかない。**

そして内訳を見ると、日本の銀行にとって笑えない偏りがある。

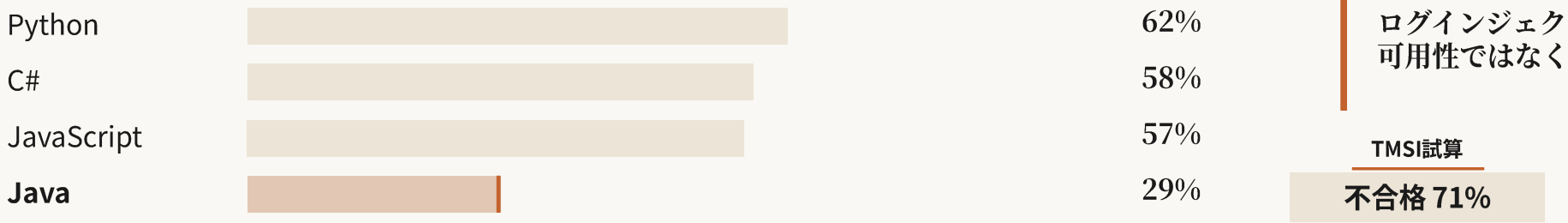
Veracode「2025 GenAI Code Security Report」（2025年7月30日）および「Spring 2026 GenAI Code Security Update」（2026年3月24日／150超のモデル・80課題・4言語・4CWE）。

AIがもっとも安全に書けない言語が、日本の勘定系の移行先だった。

同じ調査を言語別に見ると、合格率はPython 62%、C# 58%、JavaScript 57%に対し、Javaは29%。当社試算では71%が不合格ということになる。脆弱性の種類別では、SQLインジェクション82%・不安全な暗号処理86%と高い一方、クロスサイトスクリプティングは15%、ログインジェクションは13%にとどまる。推論モデルを使っても、全体の合格率は70〜72%である。



言語別の合格率

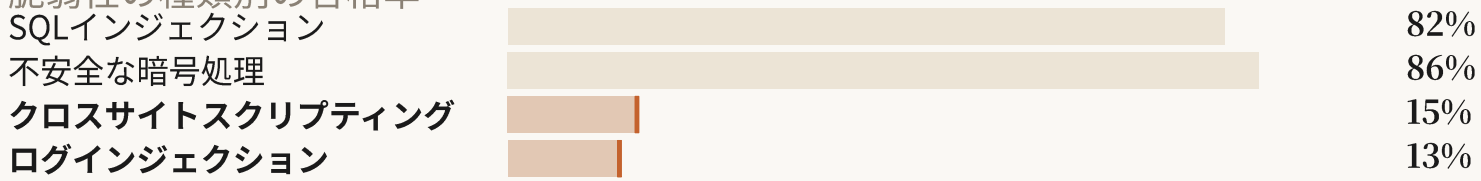


ログインジェクションの合格率13%は、可用性ではなく、統制の問題である。

TMSI試算

不合格 71%

脆弱性の種類別の合格率



推論モデルを使っても、全体の合格率は70〜72%にとどまる

日本の勘定系モダナイズが多く選ぶ移行先言語が、AIがもっとも安全に書けない言語である

監査証跡そのものが汚染されうるということだ。金融機関にとって、これは性能の議論ではない。

Veracode「Spring 2026 GenAI Code Security Update」（2026年3月24日）。不合格71%は当社試算（100－29）。CWE-80（クロスサイトスクリプティング）、CWE-117（ログインジェクション）。

速くなったのは、
書くことだけだった。

第二幕で確認したことを整理する。AIの利用は9割に達し、8割超が効果を実感している。だが実測では乖離が生じることがあり、生成物の安全性は2年間動いていない。そしてAIの導入はデリバリーの安定性に負の関係を持つ。

これらは矛盾していない。すべて「書く速度は上がったが、それ以外は上がっていない、あるいは負荷が増えた」で説明できる。

銀行のレガシー刷新において、書く速度がボトルネックだったことは、実はほとんどない。要件が決まり、仕様が確定し、影響範囲が判明していれば、コードを書く工程はもともと見積もれる工程だった。詰まっていたのは、その手前と、その後ろである。



**増幅器は、ボトルネックを解消しない。
ボトルネックの手前に、在庫を積む。**

では、詰まっているのはどこか。第三幕で、構造を分解する。

ボトルネックは、三層に分かれている。

書く。読む。確かめる。この三層で見ると、いま何が起きているかがはっきりする。第1層はAIで大きく速くなった。第2層は効き始めている——経済産業省・IPAの総括レポートも、生成AIを現行仕様の解明や設計・開発作業に位置づけた。問題は第3層である。ここが最大のボトルネックであり、しかもAIによって負荷が増える側にある。

※ 概念図であり実数を含みません

			現在の投資配分（概念）
1	書く	コードを書く。テストを書く。ドキュメントを書く	大きく速くなった
2	読む	既存システムの仕様を理解し、何をしているかを復元する	効き始めている
3	確かめる	直したものが正しいと保証し、説明できる状態にする	むしろ負荷が増える

第1層だけを速くすると、全体は速くならない。

速くした分だけ、
確かめる工程に積み上がる

現在の投資の多くは第1層に集中している。だが第1層を速くすると第3層の負荷が増えるため、全体のスループットは変わらないか、安定性の分だけ悪化する。

工程構造は当社による整理です。第3層で負荷が増える傾向については、DORA「State of AI-assisted Software Development 2025」（Google Cloud、2025年9月23日）がAI導入とデリバリー安定性の負の関係を報告しています。第2層への生成AI適用の位置づけは、経済産業省・情報処理推進機構（IPA）「DXの現在地とレガシーシステム脱却に向けて」（2025年5月28日）によります。

当局は、こう書いた。

第3層をさらに掘ると、その底に何があるか。金融庁のAIディスカッションペーパー（第1.1版、2026年3月）が、答えを直接書いている。同ペーパーは預金取扱金融機関を中心に130社から回答を得た調査に基づくもので、システム開発分野については、生成AIの活用が大きな成果をあげているという指摘も紹介している。そのうえで、28ページにこう記した。

> **開発等に外部リソースを活用する場合には
> そもそも知見や知財が組織内に残らず、
> 内部の人材育成が思うように進まない**



監督当局が、外部委託によって知見が組織内に残らないことを、公開文書で明記した。これは技術の問題ではなく、体制の問題である。

同じペーパーは、
金融機関側の対応も記録している。合弁会社を設立してグループ内に知見・知財が残る形の開発体制を作った例。大手金融機関を中心に、AIモデルの開発・運用・管理を内製化する動き。そして特定ベンダーへの過度な依存を回避するためオープンソース基盤を活用する方針。いずれも、知見がどこに蓄積されるかをコントロールしようとする動きである。

つまり論点は、外注するかどうかではない。外注した結果、何が自組織に残るように設計されているか、である。

金融庁「AIディスカッションペーパー（第1.1版）— 金融分野におけるAIの健全な利活用の促進に向けた初期的な論点整理」（2026年3月3日）p.17、p.28、p.31-32。調査期間2024年10月3日〜11月15日、回答130社（うち預金取扱金融機関が約40%）。

レガシーとは、古いコードのことではない。
決定の履歴が失われたシステムのことだ。

AIは、コードが何をしているかを高い精度で復元できる。これは実際に効いている領域であり、第2層への投資が有効な理由でもある。だが、なぜそうなっているかは復元できない。

なぜこの例外処理があるのか。なぜこの勘定科目だけ処理が違うのか。なぜこの上限値なのか。1997年の通達かもしれないし、2003年の事故対応かもしれないし、ある支店の運用上の事情かもしれない。コードには結果しか書かれていない。理由は、それを決めた人の頭の中と、失われた議事録の中にある。

経済産業省・IPAの総括レポートも、レガシーの実態をシステムのブラックボックス化、そして仕様や設計のドキュメントが整備されていないことと記述している。問題は年数ではなく、追跡可能性の欠落である。

そして銀行では、この「なぜ」こそが監査リスクの宿る場所である。理由の分からない仕様を、理由の分からないまま新環境に移植すれば、動くシステムは手に入る。だが、説明できないシステムになる。



コードは読める。
理由は、残っていない。

刷新とは、コードを新しくすることではない。理由を取り戻す作業である。

経済産業省・情報処理推進機構（IPA）「DXの現在地とレガシーシステム脱却に向けて」レガシーシステムモダン化委員会総括レポート（2025年5月28日）。レガシーの定義に関する記述は当社による解釈です。

設計書からコードへ。日本の勘定系で、すでに始まっている。

三菱UFJ銀行と日本IBMは2025年5月、勘定系システム開発のモダナイゼーションにおける協働を公表した。2024年に検証を終え、2025年から本格適用に入っている。公表内容から読み取れることを、そのまま三点に分けて示す。

2025年5月9日／日本IBM ニュースルーム

設計書から、生成AIでコードを生成する

設計書を入力として、生成AIによりPL/Iのコードを生成する取り組み。
第1層（書く）ではなく、
第2層（読む・仕様を扱う）から入る。
既存の設計資産を起点にする点が、
白紙からコードを書かせるアプローチと
決定的に異なる。

同上

開発環境そのものを寄せる

VSCodeとGitLab、GitLab CIを用いた開発環境を、IBMメインフレームとRed Hat OpenShiftの上に構築。公表されている狙いのひとつが、メインフレームと分散系の開発者の環境差異の解消である。人が行き来できる状態を作ることが先に置かれている。

同上

目的は速度ではなく、連続性

もうひとつの狙いとして挙げられているのは、マルチプラットフォームに精通した次世代勘定系エンジニアの育成である。定量的な効果目標は公表されていない。何倍速くなるかではなく、誰が引き継げるかを設計している事例として読める。

本資料はこの取り組みを評価しない。ただ、目的の置き方が本編の主張と一致していることだけを指摘しておく。

日本アイ・ビー・エム株式会社 ニュースルーム（2025年5月9日）。本件について定量的な効果目標は公表されていません。本資料は特定ベンダー・製品を推奨するものではありません。

三つのレバー。順番を間違えないこと。

以上を踏まえて、実行可能な打ち手を三つに絞る。重要なのは並び順ではなく、着手の順序である。レバー2はレバー1より先に、あるいは同時に始めなければならない。レバー3は効果が出るまでもっとも時間がかかるため、いま始める必要がある。

※ 概念図であり実数を含みません



レバー 1

読む速度に投資する

打ち手

コード解析と仕様復元にAIを適用する。生成物は新しいコードではなく、読める設計書

効果が出るまで **3〜6か月**

注意点 復元された仕様は仮説である。検証の導線を先に設計する



レバー 2

検証を先に厚くする

打ち手

生成の導入と同時に、自動テスト・静的解析・セキュリティスキャンを強化する

効果が出るまで **導入と同時（むしろ先行）**

注意点 順番を逆にしない。生成量が増えてからでは、滞留が可視化されるだけ



レバー 3

理由を残す仕組みに変える

打ち手

決定の記録を成果物として定義する。委託契約に知見が残る条件を書く

効果が出るまで **12か月以上**

注意点 残すのは仕様ではなく判断。分量を増やすと運用が破綻する

AIに書かせる前に、理由を残す仕組みを作る。
逆にすると、速く動く、説明できないシステムが手に入る。

レバー1と2を同時に。レバー3は、いま着手しないと次の刷新でも同じことが起きる。

本図は当社による整理であり、特定の調査結果を示すものではありません。

自行はどこにいるか。そして、最初の90日で何をするか。

※ 概念図であり実数を含みません

Lv.1 個人利用

開発者が各自の判断でAIを使っている。組織として把握していない

「誰がどれだけ使っているか」という問いに答えられない

多くの銀行はここ

Lv.2 生成の標準化

コーディング支援が正式に配備され、生成量が増えている。検証は従来のまま

レビュー待ちが増え、リリースの安定性が体感で落ちている

Lv.3 理解への適用

既存システムの仕様復元にAIを適用し、有識者の検証プロセスが回っている

復元された仕様が、レビュー済みの状態で参照されている

Lv.4 履歴の設計

決定の記録が成果物として定義され、委託先も含めて運用されている

5年前の仕様変更について、いまでも理由を辿れる

コードは読める。理由は、残っていない。
ならば、数えるべきは行数ではなく、理由の分からなかった箇所の数である。

最初の90日でやること

- 対象システムを1つだけ選ぶ（刷新の緊急度ではなく、仕様が失われている度合いで選ぶ）
- 復元結果のうち、有識者にも理由が分からなかった箇所を数える（これが本当のレガシー度）
- そのシステムの仕様復元をAIで試行し、有識者が正誤を判定できる形式に落とす
- 検証工程を棚卸しし、生成量が2倍になった場合に耐えられるかを見積もる

最初の90日でやらないこと

- 全社的なAI開発ツールの一斉展開（Lv.2を加速させるだけになる）
- 生産性の定量目標の設定（この段階の自己申告値は当てにならない）
- 移行先アーキテクチャの決定（現行仕様が分からないと決められない。順番が逆）
- ドキュメント整備プロジェクトの立ち上げ（分量を増やす方向は必ず破綻する）

多くの銀行はLv.2にいる。そしてLv.2は、増幅器が弱みの側を拡大している状態である。

本表および最初の90日の整理は、当社によるものです。特定の調査結果を示すものではありません。

採用しなかった数字と、その理由。

この領域は、ベンダー発の予測値と実測値が混在して流通している分野です。本資料では、発表主体の一次資料で確認できなかった数値、および実測でない予測値をすべて図表から除外しました。何を落としたかを開示しておきます。

AI活用による開発生産性3〜4倍	発表主体による見込み値であり、実測値ではない。本資料は予測値を数値として掲載しない方針のため不採用。
海外投資銀行の開発部隊約1.2万人／コーディングエージェント数千体規模への拡大計画	当該社の公式資料で該当する記載を確認できなかった。また「計画」であり実績ではない。不採用。
海外大手行が4万人の開発者にAI支援を配備／決算説明での定量報告	一次資料を特定できなかった。報道ベースでは同社のコーディングアシスタントによる効率化に関する言及が確認できるが、図表には用いない方針のため不採用。
4カ月で変換率97%のCOBOL移行	業界紙の報道でのみ確認。分母・対象範囲・手戻りの扱いが不明で、再現条件を判断できないため不採用。
地域銀行の脱COBOL完了時期	報道でのみ確認。当該行の公式サイトに該当するリリースがないことを確認したため、時期を数値として掲載していない。
メインフレーム事業の販売終了・保守終了年度	報道ベース。ベンダーの公式発表として確認できなかったため、年限を数値として掲載せず、構造的な論点としてのみ記述した（p.7）。
生成AI活用状況の日米独比較	参照した調査の本文に数値の記載がなく、掲載図からの読み取り値しか得られなかったため不採用。人材の不足感については本文に数値記載があるため採用している（p.6）。

落とした数字のほうが、残した数字より多い。予測と実測が混ざりやすい領域だからこそ、線を引いています。

出典一覧

本資料に数値として掲載しているのは、発表主体の一次資料で確認できたA級の情報のみです。「当社試算」と明記した5点は、A級数値からの単純計算です。

問うべきは「AIで開発を速くするか」ではない。
自行はいま、5年前の仕様変更の理由を、辿れるか。
銀行編（全5編）完

政府・独立行政法人

経済産業省・情報処理推進機構（IPA）「DXの現在地とレガシーシステム脱却に向けて」レガシーシステムモダン化委員会総括レポート2025年5月28日
情報処理推進機構（IPA）「DX動向2025」2025年6月26日（調査期間2025年2月上旬〜3月下旬、日本・米国・ドイツ）

当局 — 金融庁

金融庁「金融分野におけるITレジリエンスに関する分析レポート」2025年6月30日
金融庁「AIディスカッションペーパー（第1.1版）— 金融分野におけるAIの健全な利活用の促進に向けた初期的な論点整理」2026年3月3日

業界団体

一般社団法人日本情報システム・ユーザー協会（JUAS）「企業IT動向調査2025」2025年4月（配布4,500社・有効回答981社・回収率21.8%）

調査・研究

DORA「State of AI-assisted Software Development 2025」Google Cloud 2025年9月23日
METR「Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity」2025年7月10日（arXiv:2507.09089）
METR「We are Changing our Developer Productivity Experiment Design」2026年2月24日
Veracode「2025 GenAI Code Security Report」2025年7月30日
Veracode「Spring 2026 GenAI Code Security Update」2026年3月24日

企業の公式発表

日本アイ・ビー・エム株式会社ニュースルーム（三菱UFJ銀行との勘定系システム開発モダナイゼーション協働）2025年5月9日

本資料は、上記の公表資料をもとに株式会社TMSI Partnersが独自に整理・分析したものです（2026年7月時点）。「当社試算」と明記した数値は、A級一次情報からの単純計算であり、発表主体による公表値ではありません。海外の調査結果は各国・各組織の環境における測定値であり、日本の金融機関での再現を保証するものではありません。報道ベースでのみ確認できた数値、および実測でない予測値は、意図的に図表に掲載していません（p.24に一覧）。

自社への適用を相談する — https://tmsi-partners.co.jp/contact.html
株式会社TMSI Partners ／ contact@tmsi-partners.co.jp

進め方（方法論）を読む — https://tmsi-partners.co.jp/services.html

Insightsに戻る — https://tmsi-partners.co.jp/insights.html

© 2026 株式会社TMSI Partners. All rights reserved. — TMSI Partners Co., Ltd. Financial × AI Transformation