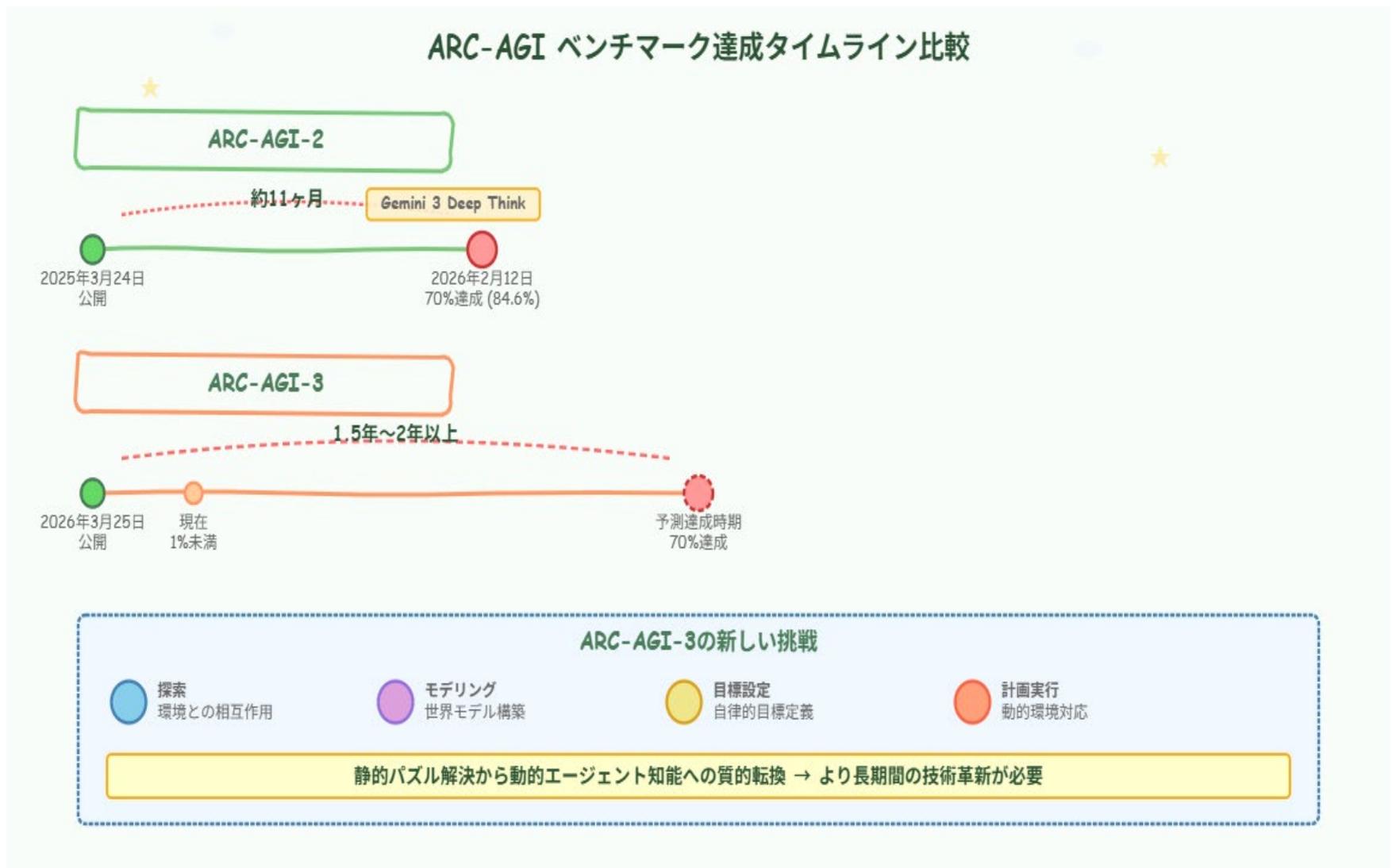


ARC-AGI-3 が出てから 70%以上の数値に達するまでの時間はどの程度と考えられるか？

Felo AI



概要

ARC-AGI-2 ベンチマークが 2025 年 3 月 24 日に公開されてから、AI が 70%以上の性能を達成するまでにかかった期間は約 11 ヶ月です [4 7 31](#)。Google の Gemini 3 Deep Think が 2026 年 2 月 12 日に正答率 84.6%を記録し、このマイルストーンを初めて達成しました [4 27](#)。

一方、2026 年 3 月 25 日に公開されたばかりの ARC-AGI-3 は、静的なパズル解決からインタラクティブな環境でのエージェント的知能の測定へとパラダイムを大きく転換させており、AI にとって質的により困難な課題となっています [6 8 36](#)。現在の AI の性能は 1%未満と極めて低く、最強クラスのモデルでもほぼ解けない状況です [18 36](#)。この質的な飛躍を考慮すると、ARC-AGI-3 で 70%以上のスコアを達成するには、ARC-AGI-2 よりも長い期間、少なくとも 1.5 年から 2 年以上を要すると予測されます。単純なスケールアップだけでは解決できず、探索、自己目標設定、継続的学習といったエージェントアーキテクチャに関する根本的な技術革新が必要となるでしょう [10 13](#)。

詳細レポート

ARC-AGI-2 : 約 11 ヶ月での 70%達成タイムライン

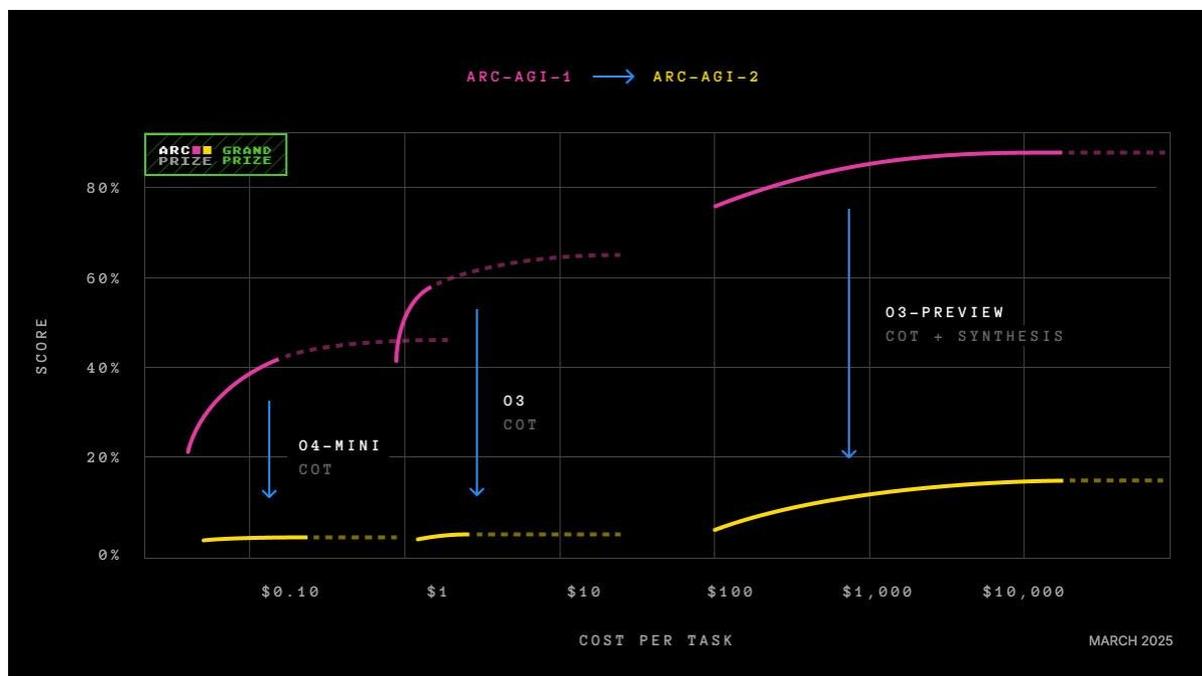
公開からマイルストーン達成まで Abstraction and Reasoning Corpus for Artificial General Intelligence 2 (ARC-AGI-2)は、AI の真の推論能力を測るために 2025 年 3 月 24 日に公開されました [7 31](#)。このベンチマークは、人間には容易だが AI には難しい抽象的なパズルで構成されており、従来の AI 評価指標とは一線を画すものとして注目を集めました [3 9](#)。

公開後、多くの主要な AI ラボがこの課題に取り組みましたが、スコアは伸び悩みました [43](#)。しかし、公開から約 11 ヶ月後の 2026 年 2 月 12 日、Google がリリースした論理推論モデル「Gemini 3 Deep Think」が、過去最高となる ****84.6%****の正答率を記録し、初めて 70%の壁を突破しました [4 27](#)。この成果は、AI の推論能力が新たな段階に入ったことを示す重要な出来事となりました。

その後、2026 年 2 月 20 日には「Gemini 3.1 Pro」も 77.1%のスコアを達成し、高性能な推論能力がより広範なモデルで利用可能になりつつあることを示しました [23 45](#)。

イベント	日付	モデル/主体	スコア	備考
ARC-AGI-2 公開	2025 年 3 月 24 日	ARC Prize Foundation	-	新しい推論ベンチマークとして発表 7 31
70%超えを初達成	2026 年 2 月 12 日	Google Gemini 3 Deep Think	84.6%	公開から約 11 ヶ月で達成 4 27

イベント	日付	モデル/主体	スコア	備考
追従モデルの発表	2026年2月20日	Google Gemini 3.1 Pro	77.1%	高度な推論能力の一般提供化 23 45



成功の要因 ARC-AGI-2 での急速な進歩は、単なるモデルの規模拡大（スケーリング）だけではなく、「リファインメントループ」や「テスト時適応」といった新しいアルゴリズムやアプローチの導入が鍵となりました [25](#) [33](#)。これらの手法は、AI が未知の問題に対してより柔軟かつ効率的に解決策を探索・改良する能力を高め、静的な知識の記憶だけでは解けない ARC-AGI の課題攻略に貢献しました。

ARC-AGI-3 : 次なる挑戦と 70%達成への道筋

パラダイムシフト：静的パズルから動的エージェントへ 2026年3月25日に公開された ARC-AGI-3 は、AI 研究の次なるフロンティアを提示しています [36](#) [41](#)。ARC-AGI-1 および 2 が静的なグリッドパズルにおける「受動的な流動性知能」を測定したのに対し、ARC-AGI-3 は AI エージェントが動的な環境内で自律的に行動し、学習する能力を評価する、世界初の「インタラクティブな推論ベンチマーク」です [6](#) [10](#) [48](#)。



この新しいベンチマークでは、AI エージェントは以下の 4 つのコア能力を試されます [10](#) :

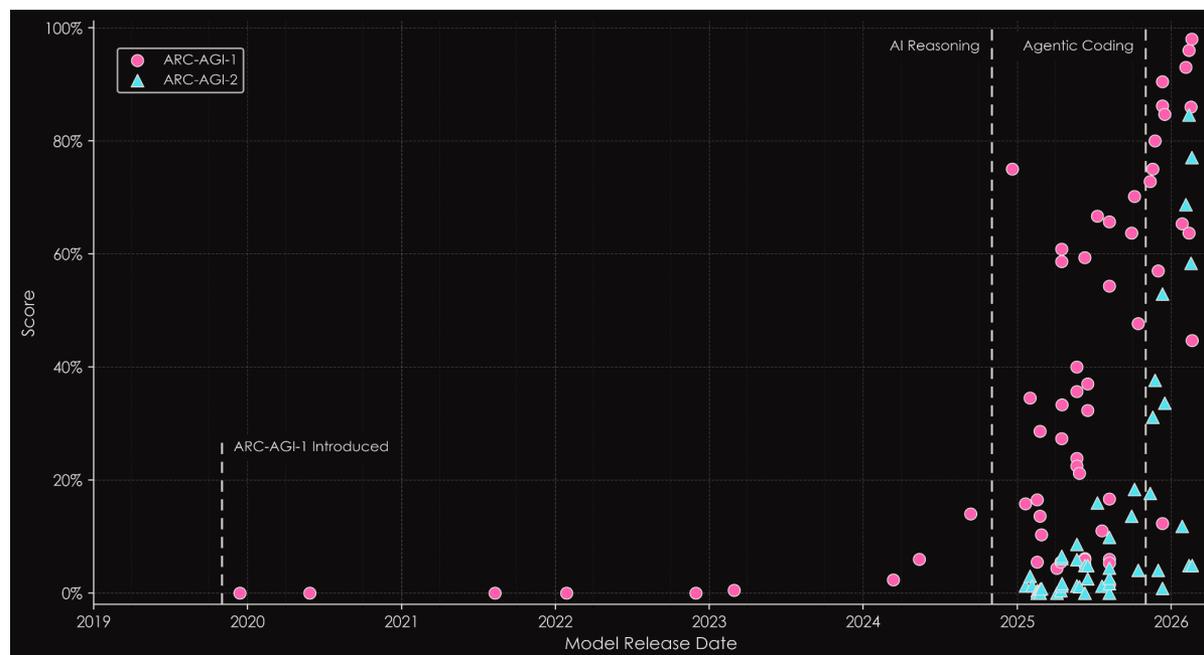
1. **探索 (Exploration):** 周囲の環境と相互作用することで、能動的に情報を獲得する能力。

2. **モデリング (Modeling):** 観測データから一般化可能な世界のモデルを構築し、未来を予測する能力。
3. **目標設定 (Goal-setting):** 明示的な指示なしに、環境からのフィードバックに基づき自律的に目標を定める能力。
4. **計画と実行 (Planning and Execution):** 現在の状態から目標達成までの行動計画を立て、予期せぬ事態に対応しながら実行する能力。

現状の性能と将来予測 ARC-AGI-3 の公開時点での AI の性能は極めて低く、**1%未満**と報告されています [18](#)。最強クラスの推論モデルでさえ、この新しい形式の課題はほとんど解けない状況です [36](#)。これは、ARC-AGI-3 が既存の AI アーキテクチャでは対応が困難な、根本的に新しい能力を要求していることを示唆しています。

ARC-AGI-2 が約 11 ヶ月で 70%の壁を越えたのに対し、ARC-AGI-3 で同様のマイルストーンを達成するには、はるかに長い時間が必要だと考えられます。その理由は以下の通りです。

- **質的な難易度の飛躍:** ARC-AGI-3 は、単に問題を解くだけでなく、「ルールそのものを発見し、適応する」プロセスを評価します [32](#)。これは、受動的なパターン認識から能動的な世界との相互作用へと、知能の定義そのものを拡張するものです。
- **必要な技術革新:** スケーリング則だけではこの課題は解決不可能と見られています [13](#)。自律的な探索戦略、継続的なオンライン学習、内発的動機付けに基づく目標設定など、エージェントアーキテクチャに関する基礎的なブレークスルーが不可欠です。
- **ベンチマークの設計思想:** ARC-AGI シリーズは、AI の能力が飽和しないように意図的に設計されており、AI の進歩に応じてベンチマーク自体も進化していきます [25](#)。ARC-AGI-3 は、AGI (汎用人工知能) への「北極星」として、今後数年間の AI 研究の方向性を導く役割を担うことが期待されています [13](#)。



これらの要因を総合的に勘案すると、ARC-AGI-3で70%以上のスコアを達成するには、ARC-AGI-2の攻略にかかった期間を大幅に上回る、少なくとも1.5年から2年、あるいはそれ以上の期間を要する可能性が高いと結論付けられます。この挑戦は、次世代のAIモデルの開発を促し、真の汎用知能への道を切り拓くための重要な試金石となるでしょう。

1. [2025 Competition Details – ARC Prize](#)
2. [ARC–AGI で新しい SOTA が達成されました : r/singularity – Reddit](#)
3. [「人間には簡単だが AI には難しい」とされるゲーム課題を実際に無料 ...](#)
4. [Gemini 新版で成績急進、「ARC–AGI」ベンチが注目される理由](#)
5. [ARC Prize 2026 – ARC–AGI–2 | Kaggle](#)
6. [ARC–AGI–3](#)
7. [うさぎでもわかる ARC–AGI–2 – 汎用人工知能への新たな挑戦 – note](#)
8. [ARC–AGI ベンチマークが注目される理由](#)
9. [Announcing ARC–AGI–2 and ARC Prize 2025](#)
10. [ARC Prize 2026 – ARC–AGI–3 Competition](#)
11. [ARC–AGI–2 と ARC 賞 2025 発表！ – Reddit](#)
12. [/keidai–reds – WWW3](#)
13. [ARC Prize](#)
14. [The ARC of Progress towards AGI: A Living Survey of Abstraction ...](#)
15. [ARC–AGI–2 概要（詳細プレゼンテーション） : r/singularity – Reddit](#)
16. [双腕協働移動ロボットによる AMR ソリューション](#)
17. [r/agi on Reddit: GPT–5 Pro just broke 70% on ARC–AGI–1 18 % on ...](#)
18. [Announcing ARC–AGI–3 The only unsaturated agentic intelligence ...](#)
19. [AGI もこうなる。新モデルが ARC–AGI–2 ベンチマークあたりを ...](#)
20. [ARC AGI 3 is up ! Just dropped minutes ago : r/accelerate – Reddit](#)
21. [ARC Prize \(@arcprize\) / Posts / X – Twitter](#)
22. [ARC Prize 2026 – ARC–AGI–3 | Kaggle](#)
23. [Gemini 3.1 Pro、ARC–AGI–2 で 77.1% “Deep Think 級推論”を一般 ...](#)
24. [ARC–AGI 3 が紹介されました – Reddit](#)
25. [ARC Prize 2025 Results and Analysis](#)
26. [Will any model score 70+% on ARC AGI 2 within 3 ... – Manifold](#)
27. [Gemini の能力は人間に近づいたか、ARC–AGI ベンチマークが注目 ...](#)
28. [ARC–AGI のベンチマークがなぜ重要なのか、そして AI にはまだどれ ...](#)

29. [What is ARC-AGI? - ARC Prize](#)
30. [ARC Prize 2026 - ARC-AGI-3 | Kaggle](#)
31. [AIの「本当の賢さ」を測る：新テスト「ARC-AGI-2」が暴く推論...](#)
32. [【対話ログ】真の汎用知能を測るベンチマーク「ARC-AGI-3」公開](#)
33. [ARC-AGI-2](#)
34. [ARC-AGI-3 Preview: 30-Day Learnings](#)
35. [主要な AI モデルが AGI テストで全滅：汎用人工知能の高い壁 \(2/3\)](#)
36. [「未知の問題を解けるのが AGI」、AI 研究者のフランソワ・ショレ氏 ...](#)
37. [arcprize/ARC-AGI-2 - GitHub](#)
38. [Interactive Reasoning Benchmarks | ARC-AGI-3 Preview : r/accelerate](#)
39. [ARC プライズ 2025 および ARC-AGI-2 の立ち上げ - LinkedIn](#)
40. [【完全ガイド】ARC-AGI とは？～概要と始め方～ - Zenn](#)
41. [ARC Prize 2026](#)
42. [Analyzing o3 and o4-mini with ARC-AGI](#)
43. [主要な AI モデルが AGI テストで全滅：汎用人工知能の高い壁](#)
44. [ARC-AGI-3 は、人間と AI の両方で高い IQ \(145 以上\) を確実に測定 ...](#)
45. [Gemini 3.1 Pro を発表 - ARC-AGI-2 で前世代比 2 倍超の推論性能を達成](#)
46. [ARC Prize 2025: Technical Report - arXiv](#)
47. [Gemini 3 は、ARC-AGI-2 ベンチマークで素晴らしい結果を達成しま ...](#)
48. [Leaderboard - ARC Prize](#)
49. [ARC-AGI-2 human baseline surpassed \(updated\) - LessWrong](#)