「ScientistTwo: Pioneering the Human Knowledge Frontier with Autonomous AI」

調査・批評レポート

― 自律型 AI 研究エージェントの主張と実証の乖離 ―

2026 年 9 月 20 日時点

Claude Fable 5.1

要旨

- **「ScientistTwo」は実在する。**arXiv:2609.19644（v1、2026 年 9 月 17 日投稿）、Google Cloud AI Research の Jaehyun Nam ら 7 名による技術論文であり、機械学習（ML）研究の 107 問中 86 問（80.4%）で人間の SOTA（最高性能）を上回り（平均相対改善 25.2%）、生成論文が自動査読で ICLR／NeurIPS 採択論文の平均点を超えたと主張する自律型 AI 研究エージェントである[1,2]。
- **主張は技術的には注目に値するが、「人類の知のフロンティア開拓」という表題は証拠に対して過大である（分析）。**評価は主に自身のループ内 AI 査読器（ScholarPeer）に依存し、独立系査読器では優位が縮小・逆転する兆候がある[5]。人間専門家による査読・独立再現はゼロである。成果はほぼ既存論文のインクリメンタル改良で、中央値の改善は 7.7%にとどまる[5]。
- **反響は現時点でほぼ皆無（投稿 3 日後）。**確認できる「批評」はすべて AI 自動生成のものであり[5,6]、名前の付いた人間研究者による評価・報道は見当たらない。コード・データ・生成論文一式の公開も未確認である（プロジェクトサイトで PDF 閲覧のみ）[2]。

【記述区分】本稿では、一次ソース（原論文・公式ページ）で確認できた事項を「確認された事実」、二次情報・機械生成レビュー経由の事項を「二次情報」、筆者の評価を「分析・推論」として区別する。URL や全文を直接確認できなかったものは参考文献欄に［未確認］等と付記した。

主要な確認事項

1. **書誌情報は確認済み。**arXiv:2609.19644、cs.AI、2026 年 9 月 17 日投稿。著者は Jaehyun Nam、Jinsung Yoon、Yanzhou Pan、Yubo Wang、Rui Meng、Parthasarathy Ranganathan、Tomas Pfister（全員 Google Cloud AI Research）[1]。プロジェクトページは scientist-two.github.io[2]。
2. **前身「ScientistOne」が存在する。**arXiv:2605.26340（Rui Meng ら、Google Cloud AI Research）[3,4]。「Chain-of-Evidence（CoE）」で全主張を証拠に紐づける検証重視システムであり、ScientistTwo はその検証枠組み（CoE Integrity Audit）を継承しつつ「能動的な分析的拡張」を加えたと位置づける[1,2]。
3. **アーキテクチャ。**6 つのエージェント群による閉ループ（Idea Generator → Evaluator → Analyzer → Writer → Peer-Review → Meta-Review）。コーディングは Claude Code（Opus 4.8）を主に使用し、代替として Antigravity（Gemini 3.8 Flash）でも動作を確認している[1,2]。
4. **主要な定量結果。**107 問中 86 問で人間 SOTA 超え、平均相対改善 25.2%（ただし中央値は 7.7%[5]）。自動査読 ScholarPeer で採択率 91.9%（平均 7.5/10）、Stanford Agentic Reviewer で 72.1%（平均 5.7/10）。全ベースライン（AI-Researcher、CycleResearcher、AI Scientist-v2、Zochi、DeepScientist、ScientistOne 等）は Stanford 系で採択率 0%[1,2]。
5. **コスト・時間。**1 論文あたり平均 2.5 日[1]、約 3,800 ドル（金額は機械生成の Pith レビューが言及する二次情報）[6]。
6. **反響はほぼ皆無。**投稿 3 日後の時点で、名前の付いた人間研究者の評価・批判、主要メディア報道は確認できなかった。存在する「レビュー」（pith.science、themoonlight.io、GitHub の jjakimoto/research-issues issue #1615 等）はすべて AI 自動生成である[5,6]。

1　前提の検証（確認された事実）

題名どおりの文書は**実在する**。一次ソースは arXiv（abs/2609.19644、html/2609.19644）[1] と公式プロジェクトページ scientist-two.github.io[2] であり、両者は数値・著者・BibTeX（nam2026scientisttwo）で一致する。前身の ScientistOne（arXiv:2605.26340）も実在し、同じ Google Cloud AI Research から出ている（Rui Meng が両論文の著者）[3,4]。

類似名称との混同に注意が必要である。Sakana AI「The AI Scientist／v2」[9,10]、Google「AI co-scientist」（Gemini 2.0 ベースの仮説生成マルチエージェント）、FutureHouse／Edison Scientific「Kosmos」「Robin」[13,14]、Intology「Zochi／Locus」[11]等はいずれも別プロジェクトである。

2　内容の整理（一次ソースに基づく）

2.1　目的・位置づけ

「問題駆動型の自律研究」を掲げる。人間が研究課題を定義し、AI が SOTA ベースラインの再現、新規仮説の形成、実験、アブレーション、論文執筆、模擬査読・リバッタルまでを人間の介入なしで実行すると主張する[1,2]。

2.2　アーキテクチャ

- Idea Generator：Limitation Extractor → Limitation Verifier → Novelty Checker → 種アイデア生成。
- Evaluator：ベンチマークのサブセットで候補を選別する「subset-first」戦略。
- Analyzer：アブレーションで有効成分を特定し、仮説を鋭敏化する。
- Writer／Peer-Review Agent／Meta-Review Agent：模擬査読・リバッタルのループ。Rebuttal Agent は文章修正ではなく新規実験を実行する。
- コーディング基盤：Claude Code（Opus 4.8）主体。Antigravity（Gemini 3.8 Flash）でも動作を確認[1,2]。

2.3　主張する成果

LLM、ロボティクス、神経科学、音声、頑健性、強化学習、ゲーム理論、プライバシー、最適化、時系列など 10 超の領域で「publication-quality」の論文と検証済みコードを生成したとする。NeurIPS 2025、ICLR 2026、ICML 2026 Spotlight の問題仕様・コードをタスク化した 107 問のうち 86 問で人間 SOTA を上回り、平均相対改善は 25.2%と報告している[1,2]。

2.4　評価方法と定量結果

採点は 2 つの自動査読器による。ScholarPeer（Gemini 3 Pro ベース[15]、in-distribution＝

ScientistTwo のドラフト精緻化にも使用）と、Stanford Agentic Reviewer（paperreview.ai、held-out）[16]である。ScholarPeer では採択率 91.9%（平均 7.5±1.3）、Stanford 系では 72.1%（平均 5.7±0.6）。Spotlight レベル（ICML 2026 Spotlight 平均 6.9／6.1）には到達していない。**人間による査読・学会採択・独立再現は行われていない**[1]。

表 1　ScientistTwo の主要定量結果

指標	値	区分・備考
人間 SOTA 超えの問題数	107 問中 86 問（80.4%）	著者の自己申告。自身が再現したベースラインとの比較[1]
相対改善	平均 25.2%／中央値 7.7%	中央値は機械生成レビューの指摘による二次情報[5]
ScholarPeer 採択率	91.9%（平均 7.5±1.3）	in-distribution（ループ内でも使用）[1]
Stanford Agentic Reviewer 採択率	72.1%（平均 5.7±0.6）	held-out[1]
ベースライン各システム	Stanford 系で採択率 0%	AI Scientist-v2、Zochi、ScientistOne 等[1]
リバッタル第 2 ラウンド前後（49 論文）	ScholarPeer 79.6%→93.9%／Stanford 73.5%→69.4%	論文 Table 5 を根拠とする[5]の指摘（二次情報）
コスト・時間	1 論文あたり平均 2.5 日、約 3,800 ドル	金額は[6]経由の二次情報
整合性監査	49 論文で全監査通過、1,814 参照中ハルシネーション 0	自動・内部監査[1]

注：生成論文数は「86」と「49」（監査・ギャラリー対象）が混在しており、監査は 49 本ベースである。

2.5　検証・整合性（CoE Integrity Audit）

4 項目の監査（スコア再現、仕様順守／報酬ハッキングの有無、参照検証、method–code 整合）を行う。フルシステムは 49 論文で全監査を通過し、1,814 参照中ハルシネーションは 0 とされる。精緻化エージェントを外すと散発的な失敗が生じる[1]。

2.6　著者が述べる限界・安全性

計算コストの高さ（1 論文 2.5 日）とコーディングエージェントへの依存を限界として挙げる。安全性・公開範囲に関する詳細な記述は限定的である[1]。

3　評判・評価・反響

現時点で人間による評価はほぼ皆無である（投稿３日後）。X（旧 Twitter）、Hacker News、Reddit、LessWrong、Alignment Forum、Substack、ブログ、主要メディアのいずれにも、名前の付いた人間研究者による ScientistTwo 固有の評価・批判は確認できなかった。

流通している「批評」はすべて AI 自動生成である点に注意を要する。

- GitHub「jjakimoto/research-issues」issue #1615[5]：「Auto Research（headless coding agent）により生成」と明記。機械生成ながら最も具体的で、「ScientistTwo は自身のループ内査読器 ScholarPeer を Goodhart 化している。リバッタル第 2 ラウンドで ScholarPeer 採択率が 79.6%→93.9%に上がる一方、held-out の Stanford 査読器では同 49 論文で 73.5%→69.4%に下がる（＝ループ内スコアをループ外品質と引き換えに買っている）」と、論文自身の Table 5 を根拠に指摘する。あわせて「25.2%は自身が再現したベースライン比の自己申告であり、中央値は 7.7%」と注記している。
- pith.science[6]：機械生成レビューであり、「人間の署名レビューはまだ無い」旨を明記。「重大な工学的前進だが、出版級論文という中核主張は実際のピアレビューを反映しない AI 査読器に依存」と評価し、約 3,800 ドル／タスク・2〜3 日という高コストにも言及する。
- themoonlight.io（Moonlight AI ツール）、scholarbro.com、supergok.com[7]：いずれも自動生成の要約・アグリゲータである（themoonlight.io、scholarbro.com の個別 URL は本調査で未取得）。

独立再現・第三者検証は確認できなかった。生成論文（86 本と称するが、ギャラリー／監査は 49 本）はプロジェクトページでの PDF 閲覧が中心で、コード・データ・論文一式の一括ダウンロードや GitHub コードリポジトリは確認できていない。「released code」という表現は内部監査を指し、公開リリースではない[2]。

4　関連システムとの比較・位置づけ

2025〜2026 年の自律型 AI 科学者の系譜は次のとおりである。

- **Sakana AI「The AI Scientist v1／v2」**：v2 は 2025 年、ICLR 2025 ワークショップ（ICBINB）で AI 単独生成論文「Compositional Regularization…」が個別スコア 6、7、6

（平均 6.33）で採択され、Sakana 公式によれば「人間投稿の 55%を上回る」水準であった（ただし 43 投稿中 AI 生成 3 本のうち閾値超えは 1 本のみで、事前合意により撤回）。方法論は Nature（2026 年 3 月 26 日）に掲載された[9,10]。

- **Intology「Zochi」**：Intology 公式ブログおよび CSPaper によれば、論文「Tempest: Automatic Multi-Turn Jailbreaking of Large Language Models with Tree Search」が ACL 2025 本会議に採択され、最終メタレビュー4/5・全 ACL 論文の上位 8.2%。「A*会議本体で査読を通過した初の AI」と主張する。後継「Locus」は長期並列実験を行う[11,12]（CSPaper の URL は本調査で未取得）。
- **Google「AI co-scientist」**：Gemini 2.0 ベースのマルチエージェントで、仮説生成・議論・進化（Elo トーナメント）に特化する。
- **FutureHouse／Edison Scientific「Kosmos」**：2025 年 11 月発表のデータ駆動発見システム。Kosmos 論文（arXiv:2511.02824、Mitchener ら）の実測で、1 回の実行あたり平均 42,000 行のコード実行・1,500 論文の読解、稼働は最大 12 時間・200 エージェントロールアウト[13]。Edison Scientific 公式によれば 7 件の発見（3 件が既存プレプリント／未公開結果の再現、4 件が新規貢献）。独立採点では**レポート記述の 79.4%が「正確」**（データ主張 85.5%、文献 82.1%）とされる。これは「再現可能性」ではなく記述の正確性の指標である。1 回の実行は 200 ドル[14]。
- **Periodic Labs**：2025 年 9 月にステルス解除（Liam Fedus、Ekin Doğuş Çubuk）。a16z 主導・Felicis が初回出資の 3 億ドルのシード（評価額約 15 億ドル。Bezos、Eric Schmidt、Jeff Dean、NVIDIA NVentures、DST、Accel 等が参加）。自律実験室×AI（材料・超伝導体）であり、物理実験を伴う点で計算実験のみの ScientistTwo と対照的である[17,18]。
- **その他**：AlphaEvolve（進化的アルゴリズム探索）、DeepScientist、AutoResearchClaw 等も存在する。
- **評価研究**：CRUX「Can AI agents conduct open-ended AI research?」（arXiv:2607.27191、Kirgis ら 24 名）は、未公開の NeurIPS 2026 投稿論文の中心問題をエージェントに与え、原著者が採点する「shadow evaluation」を実施した。エージェントはエンジニアリングは完遂したが「オープンエンドな研究問題では実質的進捗を出せず」、原著者は出力を明確に却下した[8]。SoundnessBench、BadScientist 等は LLM 査読器の脆弱性を指摘している。

表 2　主な自律型 AI 科学者システムとの比較

システム	開発主体	特徴	人間による検証
ScientistTwo[1,2]	Google Cloud AI Research	6 エージェント閉ループ。107 問の SOTA 改良、CoE 整合性監査	なし（自動査読器のみ）
ScientistOne[3]	Google Cloud AI Research	Chain-of-Evidence で全主張を証拠に紐づけ（前身）	本調査では未確認
The AI Scientist v2[9,10]	Sakana AI	AI 単独生成論文、方法論は Nature 掲載	ICLR 2025 ワークショップで査読通過（事前合意で撤回）
Zochi／Locus[11,12]	Intology	ACL 2025 本会議採択、上位 8.2% と主張	本会議の人間査読を通過
Kosmos[13,14]	FutureHouse／Edison Scientific	最大 12 時間・200 ロールアウトのデータ駆動発見	独立採点で記述の 79.4% が正確
Periodic Labs[17,18]	Periodic Labs	自律実験室×AI。物理実験を伴う	本調査では未確認
CRUX[8]	評価研究（Kirgis ら）	未公開問題での shadow evaluation	原著者が採点（結果は否定的）

注：「人間による検証」欄は本調査で確認できた範囲に限る。

ScientistTwo の相対的位置（分析）。ScientistOne が「検証可能性（CoE）」を、Sakana／Zochi が「実際の人間査読通過」を売りにしたのに対し、ScientistTwo は「多数タスクでの人間 SOTA 超え＋自動査読での高評価＋整合性監査」を一体で示す点が新規である。一方で、Zochi／Sakana が到達した**実際の人間ピアレビュー通過という一線は越えていない**。CRUX の否定的知見（オープンエンド研究は不可）と ScientistTwo の肯定的主張との乖離は、評価設計の差（クローズドな「既知 SOTA の改良」対「真にオープンな問題」）に起因すると分析できる。

5　批評（分析・推論）

5.1　主張の妥当性

「人類の知のフロンティアを開拓（Pioneering the Human Knowledge Frontier）」という表題は、証拠に照らして**誇張**と評価する。理由は次の 6 点である。

1. **評価の循環依存（Goodhart 化）。**品質の主要指標である ScholarPeer が、ドラフト精緻化ループにも使われる in-distribution の査読器である。論文自身の Table 5 が、リバッタルで

ScholarPeer スコアは上がる一方、held-out 査読器では下がる（採択率 73.5%→69.4%）ことを示しており[5]、「測定対象を最適化している」典型的な兆候である。新規性・重要性の判定を **LLM 査読器に委ねている**点が最大の弱点である。

2. **人間評価の不在。**新規性・科学的意義を判断した人間専門家がいない。Sakana／Zochi が（限定的とはいえ）人間査読を通したのと対照的である[10,11,12]。
3. **成果のインクリメンタル性。**タスクは既存トップ会議論文の SOTA 改良であり、中央値の相対改善は 7.7%[5]。「新しい問題定式化や研究方向の発見」は実証されていない（機械生成レビューも同旨を指摘[6]）。
4. **評価データ汚染リスク。**ベンチマークが最近のトップ会議論文由来であり、基盤モデルの学習データと重複する懸念がある（LLM 評価一般で知られる問題）。改善が「真の発見」か「記憶の再利用」かの切り分けが弱い。
5. **自己評価バイアス・チェリーピッキング。**86/107 という成功率と平均 25.2%は、著者による自己申告であり、自身が再現したベースラインとの比較である。head-to-head の比較でない旨は論文自身が認めている[1]。失敗 21 問を含めた分布や統計的有意性の扱いには留保が必要である。
6. **再現性。**整合性監査は自動・内部のものである。公開アーティファクト（コード・データ）が確認できず、第三者による独立再現は不可能な状態にある[2]。

5.2 方法論上の強み

（a）閉ループにアブレーションとリバッタル（新規実験を伴う）を組み込んだ点、（b）CoE 整合性監査によるハルシネーション参照 0・スコア再現の担保、（c）多領域・多数タスクでの体系的評価、（d）コーディング基盤の差し替え可能性を示した点は、工学的には堅実な前進である。

5.3 弱み・未解決課題

LLM 査読器への依存、人間評価・独立再現の欠如、汚染管理の弱さ、高コスト（3,800 ドル／タスク級[6]）、オープンエンド性の欠如（CRUX の知見[8]との整合）、安全性・ガバナンス記述の薄さが挙げられる。

5.4 誇大広告（hype）リスク

表題および「autonomous pioneer capable of pushing the frontiers of human discovery」等の表現は、実証内容（既知 SOTA の自動改良＋AI 査読での高評価）を上回る。読者は「AI が人間科学者を超えた」と誤読しやすく、supergok 等の二次情報も同様の注意喚起をしている[7]。

5.5 今後の影響

- **査読・学術出版**：AI 生成論文の氾濫と LLM 査読器の Goodhart 化は、ピアレビューの信頼性を脅かす。ICLR 2026 が LLM 生成論文・査読への対応方針を出すなど、学会側の制度対応が進行中である。
- **研究者の役割**：問題定義・意義判断・オープンエンドな探索という上流が人間の役割として残る（ScientistTwo も人間が課題を定義する分業である）。
- **安全性・ガバナンス**：LLM 生成コードの自律実行（Sakana でもサンドボックス必須と指摘されている[19]）、報酬ハッキング（本論文でも 1 件検出・除外[1]）への対策が要る。

5.6 知財実務への含意（補足）

AI 自律生成の「発見」が増えれば、（a）発明者性（多くの法域で発明者は自然人に限定）、（b）大量の機械生成先行技術による新規性・進歩性判断の圧迫、（c）進歩性における「当業者」水準の再定義、といった論点が現実味を帯びる。ただし、ScientistTwo の成果は主に ML 手法のインクリメンタル改良であり、直ちに特許実務を変えるものではない。

6 提言

- **現段階の扱い。**ScientistTwo は「有望な工学デモ」として扱い、「AI が人類の知のフロンティアを開拓した」という表題は割り引いて読むべきである。生成論文をそのまま引用・信頼せず、コード・データが公開され独立再現がなされるまで結論を保留する。
- **段階的な検証ベンチマーク。**（1）著者がコード・データ・全生成論文を公開したか、（2）held-out（Stanford 等）や人間査読での評価が主指標に格上げされたか、（3）CRUX 型のオープンエンド／未公開問題での外部評価に耐えたか。これらが満たされれば評価を上方修正してよい。

- **判断を変える閾値。**①実際の人間ピアレビュー（本会議級）を人手介入なしで通過、②独立第三者による再現成功、③汚染管理済み・時間的リークのないベンチマークでの同等性能、のいずれかが確認されれば「フロンティア開拓」の主張は実証に近づく。逆に、held-out 査読での優位消失や再現失敗が示されれば、主張は棄却されるべきである。
- **導入検討者向け。**3,800 ドル／タスク・2.5 日のコストに見合うのは、明確に定義済みで SOTA 改良型の ML 問題に限られる。真にオープンな探索や高信頼が要る領域には現時点で不適である。

7　未確認事項・留意点

- **反響の不在の解釈。**投稿からわずか 3 日であり、人間の評価が今後現れる可能性は高い。「批判が無い＝良い」ではなく「まだ評価されていない」と解すべきである。
- **AI 生成レビューの扱い。**本稿で引用した pith.science[6]、GitHub issue #1615[5]等の指摘は機械生成であり、人間専門家の見解ではない。ただし、その多くは論文自身の Table 等を根拠にしており、事実確認可能な点は分析に採用した。
- **一次ソースの一部未取得。**arXiv HTML 全文[1]・Pith 全文[6]はレート制限／ボット検知により完全には取得できず、要点はスニペット等で補った。3,800 ドル／タスク、Table 5 の 73.5%→69.4%、中央値 7.7%等の数値は二次・機械生成情報経由であり、原論文での直接確認は未了である。
- **数値の内部整合。**生成論文数が「86」と「49」（監査対象）で混在している。監査は 49 本ベースである。
- **head-to-head 不成立。**25.2%等は ScientistTwo が再現したベースライン比であり、厳密な対人間の直接比較ではない旨を論文が認めている[1]。
- **コード・データ公開状況。**現時点で公開コードリポジトリ・データ・論文一括ダウンロードは未確認である。将来公開される可能性はある[2]。

参考文献

URL は調査時点（2026 年 9 月 20 日）のもの。［未確認］等の付記は、URL または全文を本調査で直接確認できていないことを示す。

[1] Nam, J., Yoon, J., Pan, Y., Wang, Y., Meng, R., Ranganathan, P., Pfister, T. "ScientistTwo: Pioneering the Human Knowledge Frontier with Autonomous AI." arXiv:2609.19644 (v1), Google Cloud AI Research, 2026 年 9 月 17 日. https://arxiv.org/abs/2609.19644 ［**HTML 版 https://arxiv.org/html/2609.19644 は全文の一部のみ取得**］

[2] ScientistTwo プロジェクトページ. Google Cloud AI Research, 2026 年. https://scientist-two.github.io/

[3] Meng, R. et al. "ScientistOne: Towards Human-Level Autonomous Research via Chain-of-Evidence." arXiv:2605.26340, 2026 年. https://arxiv.org/abs/2605.26340

[4] Hugging Face Papers. "ScientistOne: Towards Human-Level Autonomous Research via Chain-of-Evidence"（論文紹介ページ）. https://huggingface.co/papers/2605.26340

[5] "Takes: ScientistTwo's headline that its papers out-rate human-authored ones is Goodharting its own in-loop reviewer…" GitHub jjakimoto/research-issues, Issue #1615, 2026 年 9 月 19 日. https://github.com/jjakimoto/research-issues/issues/1615 ［**AI 自動生成のレビュー**］

[6] "ScientistTwo: Pioneering the Human Knowledge Frontier with Autonomous AI · Pith Review." pith.science, 2026 年. https://pith.science/paper/2609.19644 ［**AI 自動生成のレビュー／全文の一部のみ取得**］

[7] "ScientistTwo Automates Autonomous AI Research." supergok.com, 2026 年. https://supergok.com/scientisttwo/ ［**自動生成系の二次情報**］

[8] Kirgis, P. et al. "Can AI agents conduct open-ended AI research? Early evidence from two case studies." arXiv:2607.27191, 2026 年. https://arxiv.org/abs/2607.27191

[9] Yamada, Y. et al. "The AI Scientist-v2." arXiv:2504.08066, 2025 年. https://arxiv.org/abs/2504.08066

[10] Sakana AI. "The AI Scientist: Towards Fully Automated AI Research, Now Published in Nature." 2026 年. https://sakana.ai/ai-scientist-nature/

[11] Intology. "Zochi Publishes A* Paper." 2025 年. https://www.intology.ai/blog/zochi-acl

[12] Andy Zhou（Intology 共同創業者）個人サイト. https://www.andyzhou.ai/

[13] Mitchener, L. et al. "Kosmos: An AI Scientist for Autonomous Discovery." arXiv:2511.02824, 2025 年. ［**URL 未取得**］

[14] Edison Scientific. "Announcing Kosmos." 2025 年 11 月. https://edisonscientific.com/news/announcing-kosmos

[15] Goyal et al. "ScholarPeer: A Context-Aware Multi-Agent Framework for Automated Peer Review." arXiv:2601.22638, 2026 年. https://arxiv.org/abs/2601.22638

[16] "Exploring Stanford Agentic Reviewer: Equilibrium Perspectives and the Paradox of AI-Driven Peer Review." ResearchGate.

https://www.researchgate.net/publication/399854497_Exploring_Stanford_Agentic_Reviewer_Equilibrium_Perspectives_and_the_Paradox_of_AI-_Driven_Peer_Review

[17] Contrary Research. "Periodic Labs Business Breakdown." 2026 年. https://research.contrary.com/company/periodic-labs

[18] TechCrunch. Periodic Labs に関する記事, 2025 年 10 月 20 日. ［**URL 未確認**］

[19] Ry Walker Research. "AI Scientist (Sakana AI)." https://rywalker.com/research/ai-scientist