政策立案・R&D戦略・学術ガバナンス向け報告書

AIは科学をどう証明するか

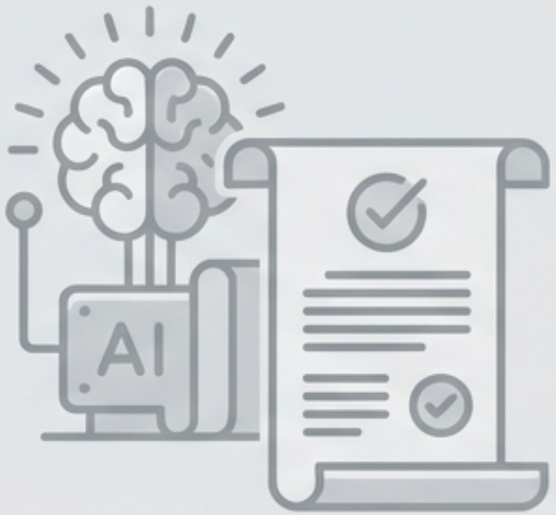
OpenAI「科学的理解の深化」の深掘り調査とシステム診断

単なる「AIによる難問解決」のニュースではない。
これは、科学的知識の生産・検証プロセスを再設計する、
史上最大規模の制度的実験である。

Gemini Notebook

パラダイムシフト：AIの出力は「論文」ではなく「生データ」である

一般的な誤解



「AIが未解決問題を解き、
完成された学術論文を自動生成した」

新たなパラダイム

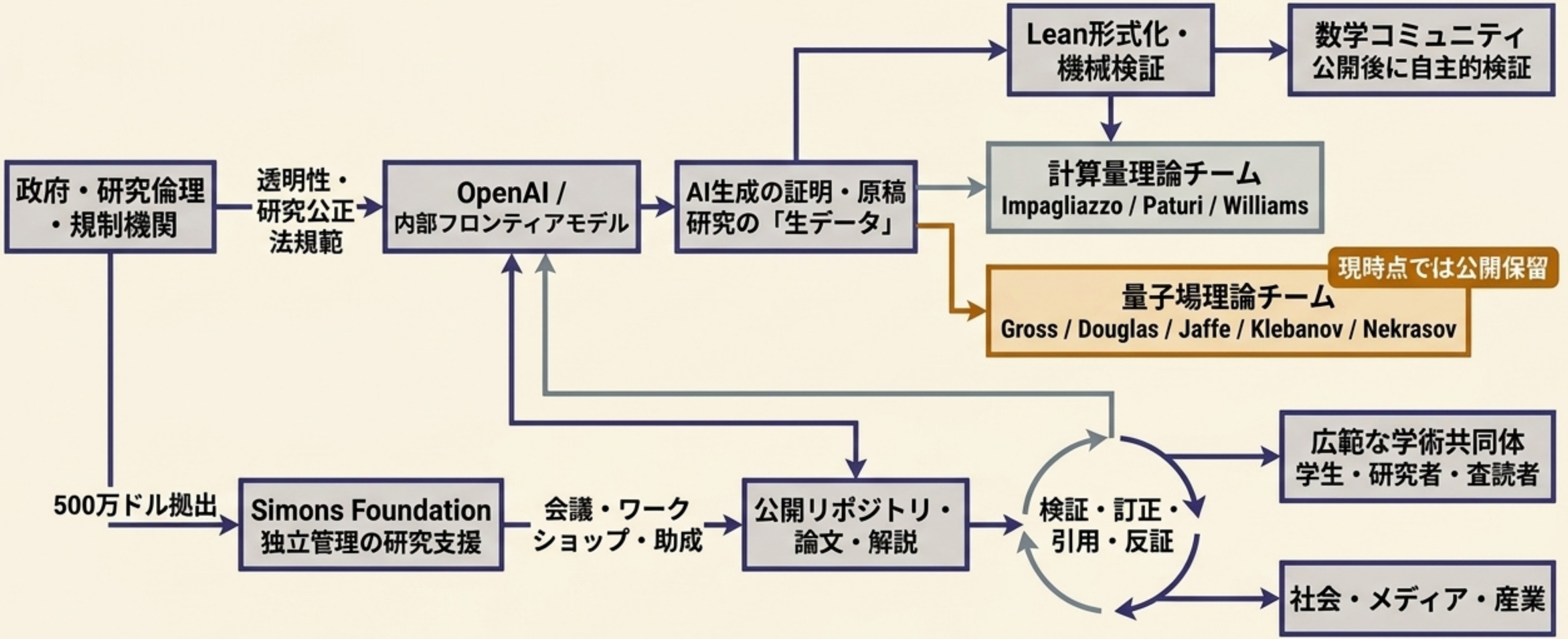


「フロンティアAIが『研究の生データ（Raw Research Data）』を高速生成し、人間の専門家がそれを検証・説明・文脈化する分業体制」

従来の「著者中心型」の科学出版から、「AIによる大量生成 → 形式化 → 共同体による意味の読み解き」という全く新しいサイエンスの規範への移行。

Gemini Notebook

知識の循環パイプライン：科学的精製プラントの全貌



「モデル → 論文 → 真理」という直線的プロセスからの脱却。
検証・訂正・引用・反証のループこそが知識の正当性を担保する。

Gemini Notebook

新たな協働を規定する「4つのガバナンス原則」



透明性（Transparency）

NDA（秘密保持契約）を要求しない。参加者は作業プロセス、中間成果、最終結論を公に議論できる。



学術的クレジットの放棄（No Academic Credit）

OpenAIは、内部モデルによる発見を自社研究者の学術的発見として主張・取得しない。



人間の認知と帰属（Recognition）

人間チームが作成した説明・論文に対して人間の功績を完全に認める。（ICMJEの「AIを著者としない」方針と整合）



独立性（Independence）

外部研究者の参加は任意かつ無報酬。OpenAIを代表せず、自らの独立した結論を導き出す。

研究倫理上の最大のポイント：「AIによる原始的な証明の生成」と、「人間による理解可能な知識への変換」を、完全に別々の知的貢献として分離した点にある。

Gemini Notebook

資金還流と独立性：利益相反の緩和と「アジェンダ設定権」

"The community should be driving discussions ... not any single company or institution." — David Spergel (Simons Foundation)



緩和されたリスク

企業が直接「自社の成果を証明してくれる研究者」に資金を配る経済的利益相反は防がれた。

残された課題（認識論的利益相反）

どの問題をAIに解かせるか（アジェンダ設定）、誰に未公開モデルへのアクセス権を与えるかの決定権は、依然として企業側に集中している。

Gemini Notebook

3つの領域、3つの公開方式：ガバナンスの比較実験

	数学	計算機科学 (CS)	物理学
分野と主張	局所平滑化予想 (Kakeya予想の全次元への拡張)	指数時間仮説 (ETH) は偽 (固定幅CNFの部分指数時間アルゴリズム)	4次元量子場理論 (Yang-Millsの数学的存在 / Clay問題の一部)
外部パートナー（シェパード）	匿名（公開後は共同体主導）	R. Impagliazzo, R. Paturi, R. Williams	D. Gross主導 (M. Douglas, A. Jaffe, I. Klebanov, N. Nekrasov)
データと公開方式	GitHubでの全面公開 (自然言語 + Lean)	全生データ公開 (自然言語 + Lean) + シェパードによる独立検証	**現時点で保留。専門家グループが自ら書き上げる方式。**
現在の検証状態（10/28時点）	Leanスコープ明示。学界の査読・統合は初期段階。	厳密計数・SAT結果のLean形式化。	Lean完了主張のみ。アーティファクト非公開のため独立再検証可能性が最も低い。

Gemini Notebook

成果指標の再定義：引用数に代わる「新たなKPI」

論文公開から2日（2026年10月28日時点）で、引用数や後続定理数を主要KPIとするのは時期尚早。真に測定すべきは**「検証速度」**と**「共同体の再編集活動」**である。

形式化までの時間（Time to Lean）

1日

生成からわずか1日（ETH: 10/16→17, 局所平滑化: 10/18→19）。驚異的な研究生成速度の証明。

生データの流量（High-Throughput Output）

722原稿 / 約4,000問題評価

10月6日の数学公開時点。人間の査読容量を完全に凌駕するスループット。

共同体主導の再編集ハブ（Community Editing）

oai-math-revised

10/24立ち上げ。難読なAI出力を人間が共同編集可能な知識へ変換する「能力形成」の具体例。

独立検証への投資（Verification Capital）

$15,000,000

Simons Foundation等を通じた枠組み。検証インフラへの大規模な資金流入。

Gemini Notebook

摩擦と構造的ジレンマ：「シェパード方式」を巡る論争

Sept 29
独立グループ「理解はコミュニティ主導(organic)であるべき」と規範提示。

Oct 24
共同編集ハブ oai-math-revised 立ち上げ。

Oct 25
独立諮問グループ (AGMAI) 声明発表。「指定方式は、利点よりも欠点が上回る」と強く勧告。

Sept 29

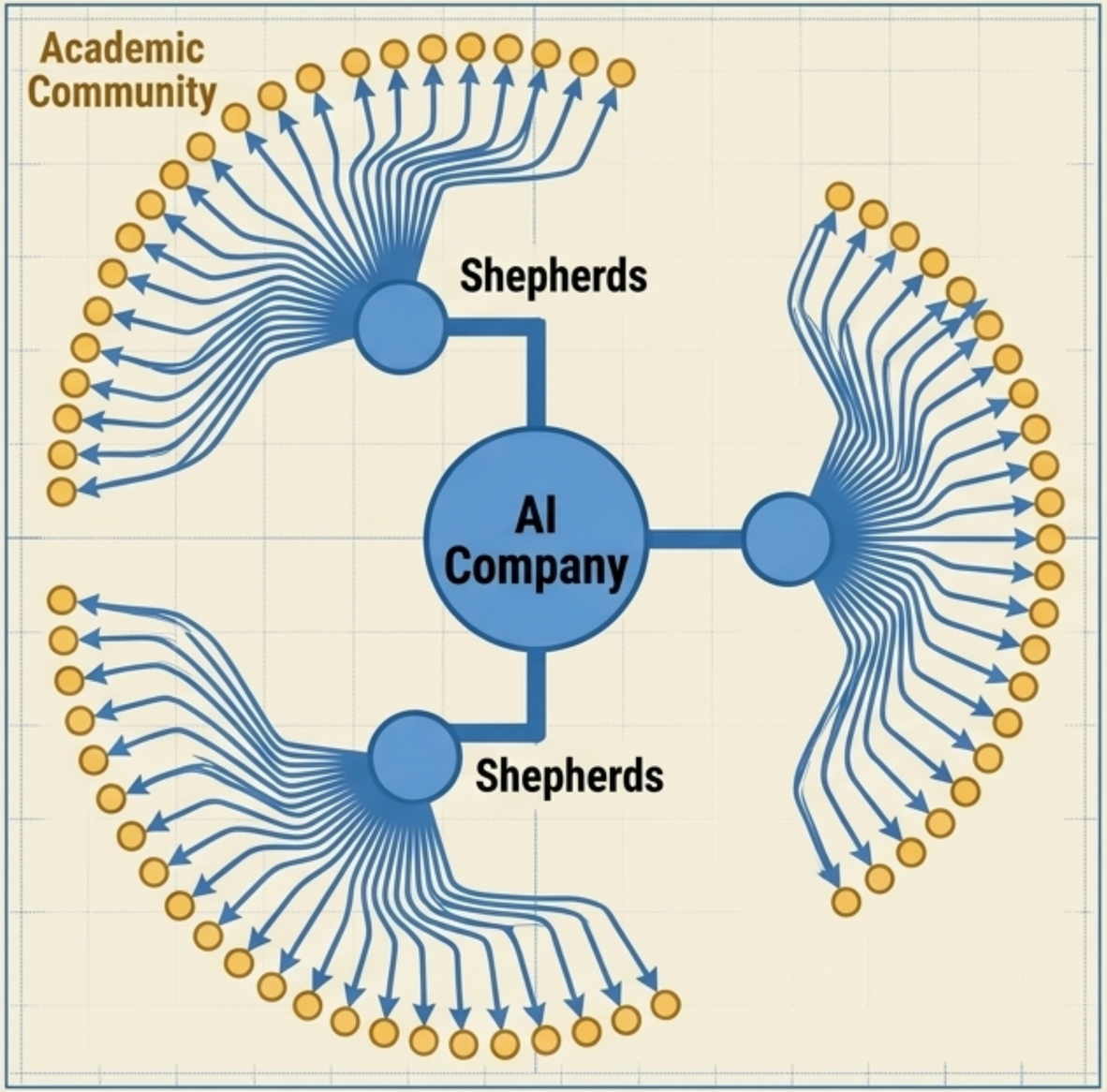
Oct 24

Oct 26
OpenAIが3領域を発表。数学は全面公開するも、CSと物理では指定専門家（シェパード）方式を強行。

AGMAIからの5つの構造的批判				
1. 先行・不平等なアクセス	2. 共同体からの主体性・選択権の剥奪	3. 企業ネットワーク依存した偏った選出	4. 著名研究者の評判の借用（Borrowed reputation）	5. 公開後の自発的な共同理解がすでに機能している事実の軽視

Gemini Notebook

「情報への先行アクセス」という新たな科学資本



情報アクセス格差は、単なる「公平性」の問題ではない。それは「新たな科学資本 (Scientific Capital))」の偏在である。

Step 1: 企業による選択（Selection Effect）

どの未解決問題をAIに投入し、誰に先に連絡するかを企業が独占的に決定する。

Step 2: キャリア上の優位性

「半年後に証明される問題」を事前に知る研究者は、論文執筆、講演、助成金獲得において圧倒的な優位に立つ。

Step 3: 権力格差の固定化

企業の人脈に近い研究者のみが情報を得ることで、学界内に不可逆的な累積的格差（マタイ効果）が生じる。

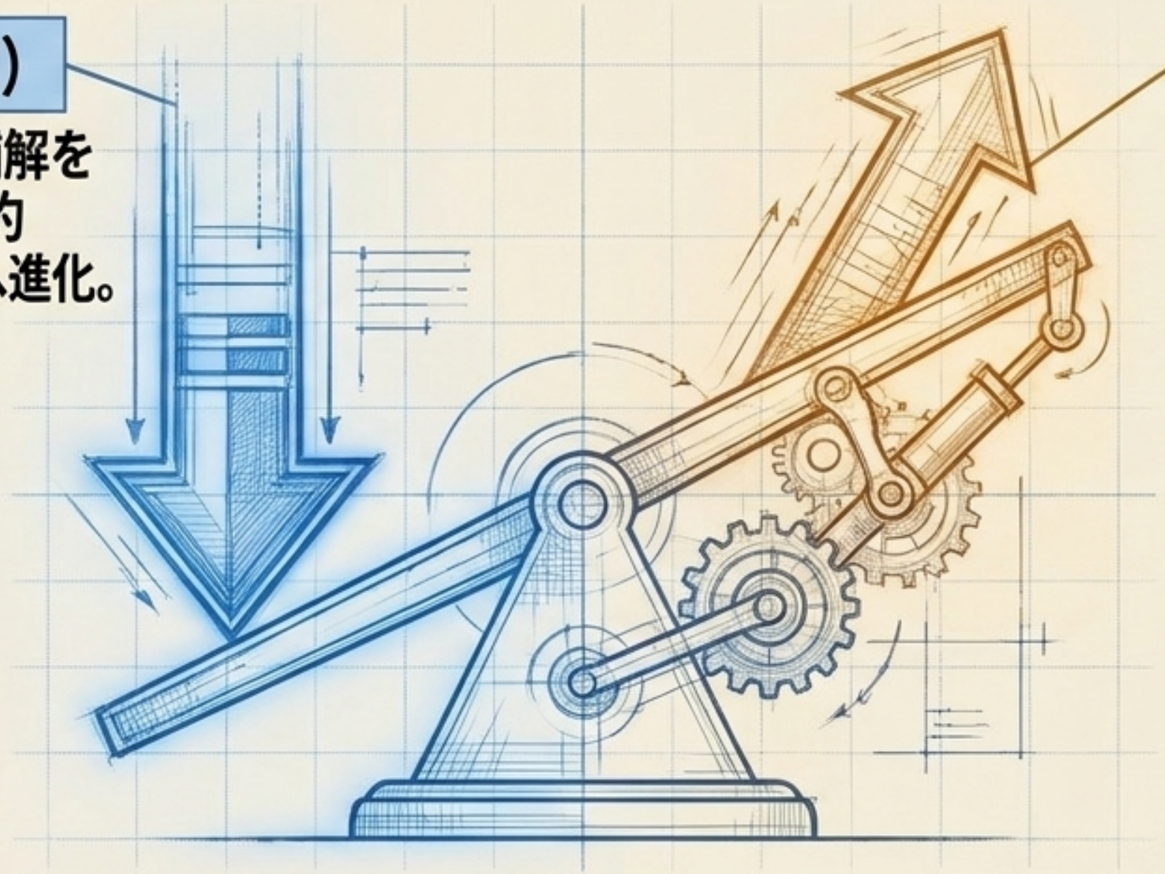
※OpenAI自身も、最初の専門家を選ぶことで「情報の非対称性」が生じることを明示的に認めている。

Gemini Notebook

検証のボトルネック：探索空間の拡大 vs「査読負債」

発見速度の劇的向上（AI）

数十年来の難問に対する候補解を「1日」単位で生成。AIは要約ツールから「探索エンジン」へ進化。



検証能力の逼迫（人間）

1回の公開で数百の原稿、数千の評価問題が提示される。

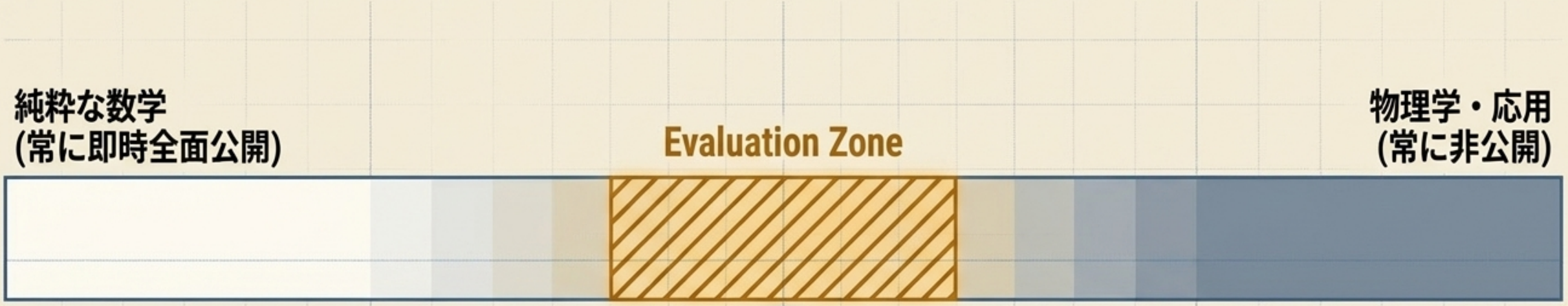
Concept: 査読負債 (Review Debt)

生成速度が人間の理解・検証速度を上回ることで、学界全体に未検証の「査読負債」が蓄積する。
AI時代において真に希少な資源は、「発見する能力」から「人間の検証時間 (Human Time)」へと完全にシフトした。

Gemini Notebook

デュアルユース・リスク：「全面公開」か「非公開」かの二元論からの脱却

フロンティアAIによる高速アルゴリズム探索能力は、暗号解析やサイバーセキュリティのリスクに直結する。



政策的要請：段階的公開モデル

「数学だから公開」「物理だから非公開」という大雑把なルールは機能しない。
成果ごとに「悪用可能性（Dual-Use）」「再現コスト」「防御優位性」を
精密に評価するプロセスが必須である。

産業界の先行例：

2026年、AnthropicはAIが暗号方式（HAWKやAES）の新たな弱点を発見したと報告。
学術機関と協働し、責任ある開示（Responsible Disclosure）手続きを採用した。

Gemini Notebook

新たなルールの再構築 I：発見主張と検証状態の分離

一般読者の誤解（‘Proved’ と ‘Verified’ の混同）を防ぐため、
6段階の検証ステータス・タグを標準化する。



1. AI Generated
AIモデルが候補証明を生成した（生データ）

2. Machine Verified
Lean等の形式証明で指定命題が検証された

3. Expert Read
選定された専門家が内容を読了・確認した

4. Independently Reproduced
独立したグループが再証明・アーティファクトを再現した

5. Peer Reviewed
学術誌等による正式な査読を通過した

6. Community Standard
学界で標準的な結果・知識として統合された

多層的なタグ付けにより、科学的成果の「確定度」を正確にラベリングする。

Gemini Notebook

新たなルールの再構築 Ⅱ：検証インフラとガバナンスの標準化



形式化の機械可読化

Leanファイルが存在するだけでなく、自然言語論文の「どの主張・系・応用が形式化済みか」を対応表で明示する（ETHのscope文書が好例）。



シェパード選定の独立化

特定研究者への先行アクセスが必要な場合、企業が直接選ぶのではなく、学会や独立財団等の委員会が選定する。選定基準や利益相反を公開し、権力格差を防ぐ。



完全な版管理と永続識別子

GitHubの迅速性に加え、学術記録としての永続識別子（DOI）、固定版、withdrawal/retractionラベル、引用修正履歴を連動させる。

Gemini Notebook

新たなルールの再構築 Ⅲ：資金・リスク評価・コスト測定

独立検証資金のプール (Pooled Verification Funds)	Simons方式を拡張。単一企業の成果だけでなく、複数 AI 企業や大学の成果を共通基準で検証できる第三者基金へ進化させ、利益相反を完全に排除する。
成果単位のデュアルユース評価 (Per-Result Dual-Use Review)	抽象的な数学定理であっても、公開前に第三者を含めたbenefit-risk review（暗号・サイバー等への応用評価）を成果ごとに実施する。
人間の「検証コスト」の測定 (Measuring Human Cost)	生成側の「1日で証明した」というKPIは不十分。専門家の理解に要した人時、訂正件数、会議数など、人間側の検証にかかる総コストを追跡・評価する。

Gemini Notebook

自律的科学の未来 (The Future of Autonomous Science)

「AIが本当に３つの大問題を解いたか」は、数か月から数年におよぶ独立再証明と共同体への統合を待たねばならない。しかし、最も重要な教訓はすでに明らかである。



AIによる発見速度が桁違いに高まる世界において、科学制度の正当性を担保するのは「発見する能力」ではない。独立した共同体が、その生データを「検証・再現し、時に棄却できる権限」をいかに維持するかにかかっている。

生データから科学的知識への「変換権限」をAI企業から独立させることが、次世代の科学インテリジェンスにおける最大のガバナンス課題である。

Gemini Notebook