# AI研究自動化と「知能爆発」：22名の専門家論文が示す事実と政策課題

2026年9月、Hinton氏、Bengio氏ら22名による論文は、AIによるAI研究自動化の急速と、知能爆発の不確実性、そして可視化と監査を中心とした低後悔型政策の必要性を報告した。

## 実証された現状（R&D自動化の急拡大）

### AIフロンティア企業における研究自動化の急激な進行



Anthropic



**26%**
R&D業務を
AIが主導

AnthropicではR&D業務の
26%をAIが主導

OpenAI



**3.1**
エージェント日 /
研究者1人1日

OpenAIでは研究者1人1日あたり
3.1エージェント日を活用

### 現状は「完全自律」ではなく人間の介入と判断が不可欠

- 4〜8時間級タスクの過半数で人間の介入が必要であり、高度な研究方針や評価は人間が担っている



## シナリオの不確実性と求められる政策

### 「知能爆発」は確実な予測ではなく条件付きシナリオ



知能爆発

**ボトルネック**

- 計算資源・データの枯渇
- 研究の収穫逓減
- 実験の物理的時間が加速を阻む摩擦



条件付きシナリオ

### 拙速な開発停止ではなく「内部R&Dの可視化と第三者監査」を実施



共通指標による
自動化率の計測



インシデント報告



低後悔型政策が
最優先課題

Gemini Notebook