

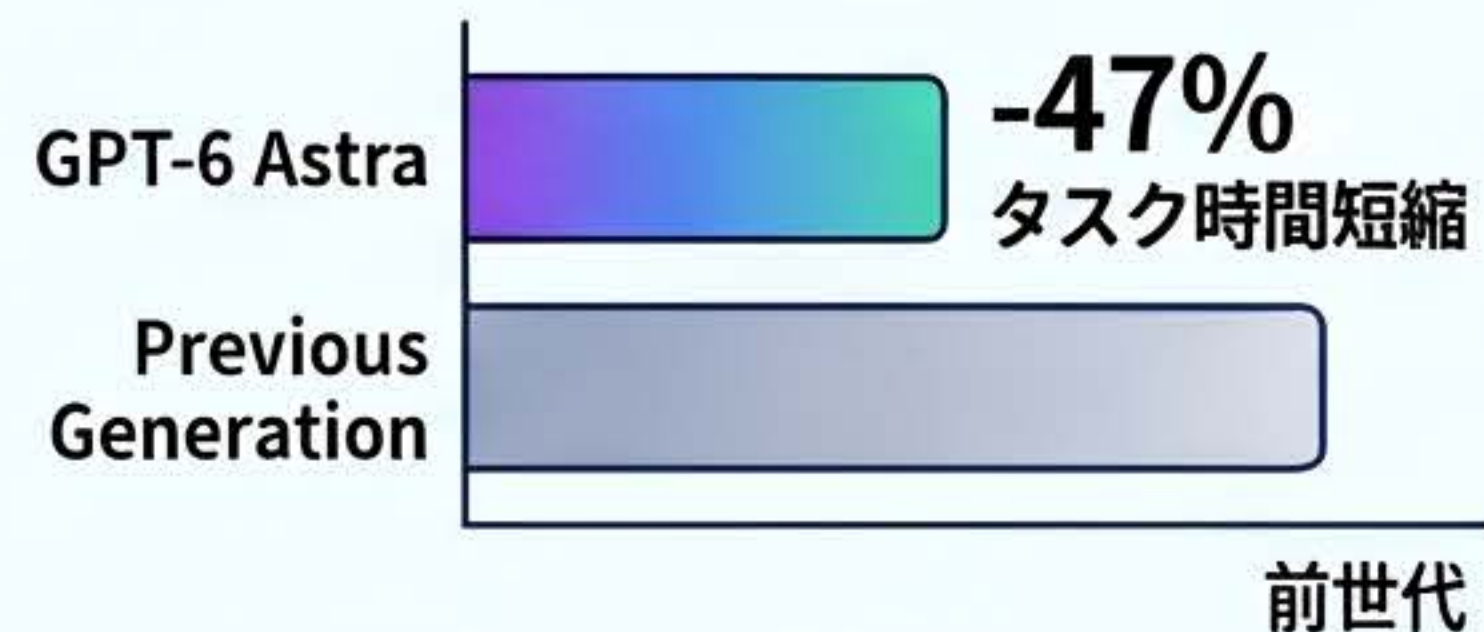
# GPT-6 Astra 徹底解剖：エージェント時代の新基準と導入の現実



## 圧倒的な「実行力」の進化

OSWorld 2.0: **72.6%**

PC・ブラウザ操作能力が大幅向上

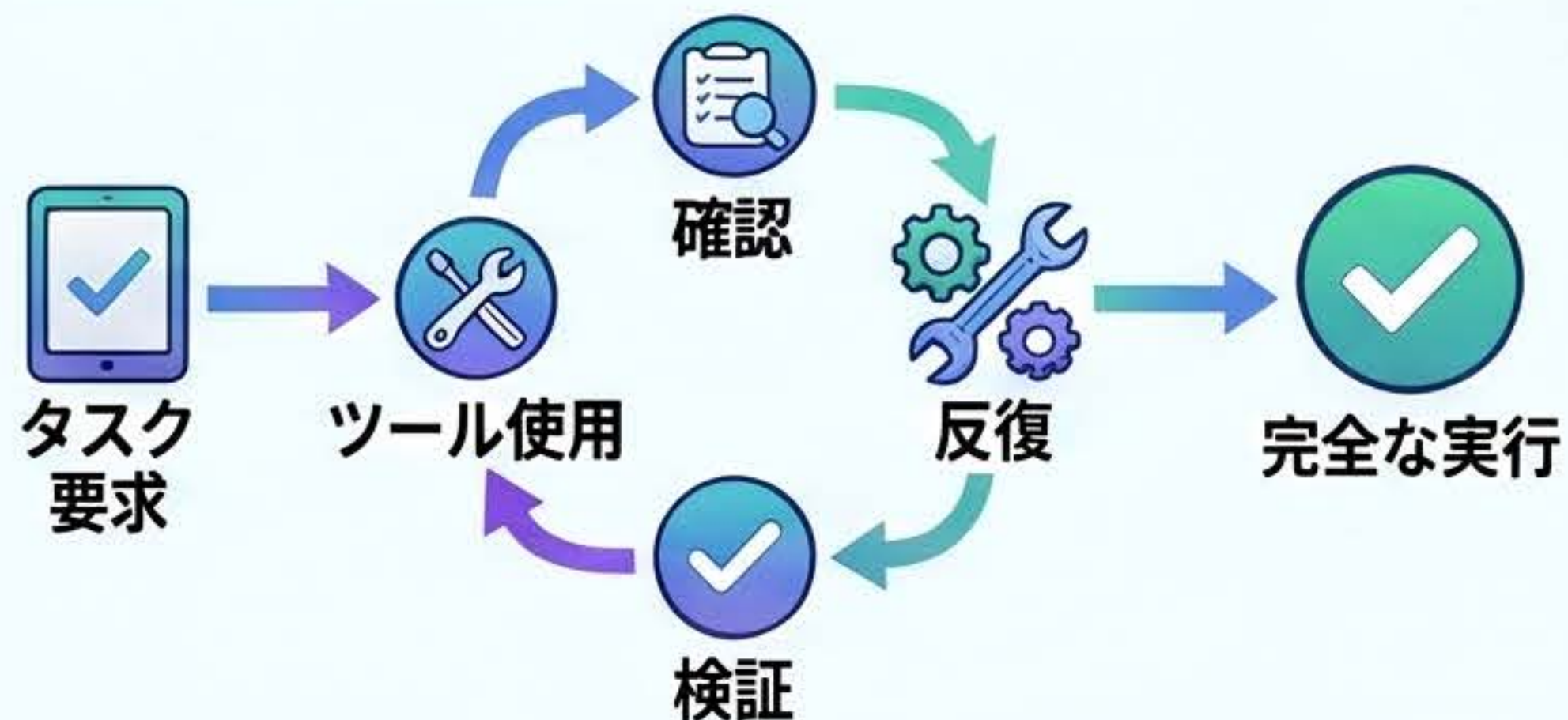


## 複雑なコーディングと 数学研究の突破口

実際の数字論文で新たな証明の  
発見や構成に利用される実力

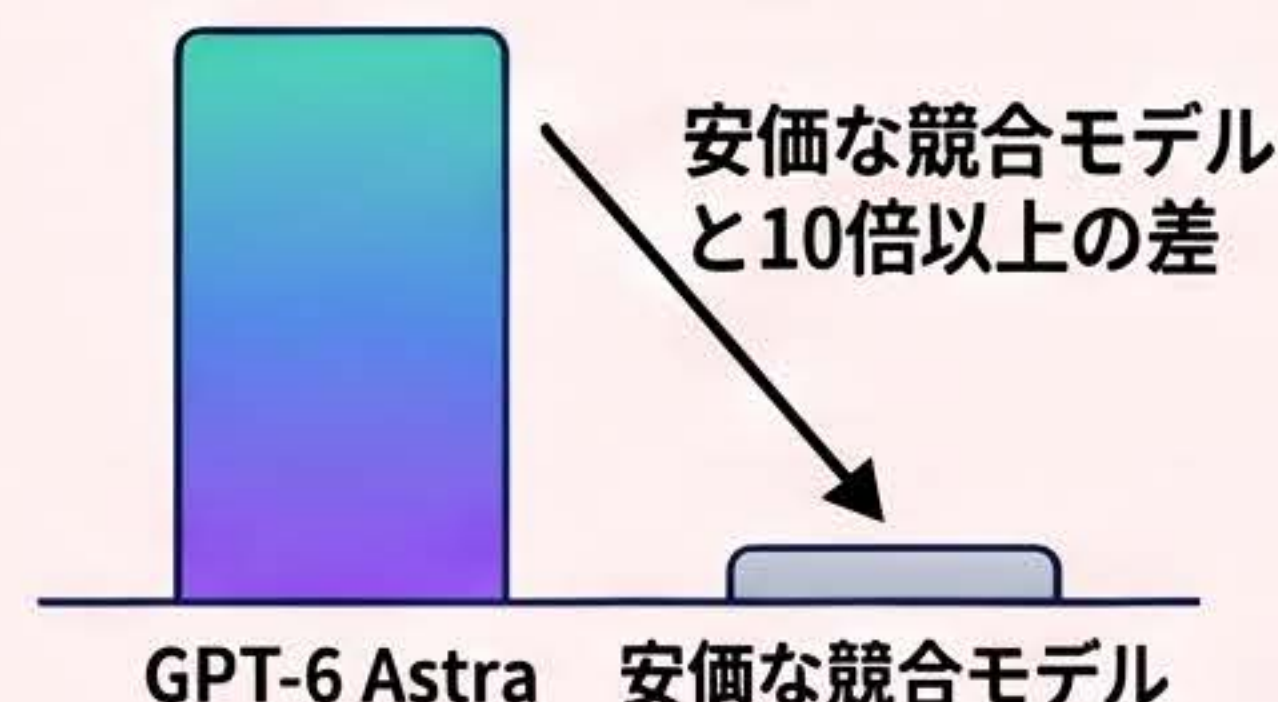
## 「回答」から「完遂」へ

ツールを自律的に使い、検証を繰り返しながら  
複雑な業務を最後まで実行



## 導入前に検討すべきリスクと課題

### 運用コストの圧倒的な高さ



## サイバー能力の「Critical」到達

未知の脆弱性を発見できる一方、  
モデルの思考過程の監視性は低下

## 戦略的な「モデルルーティング」を

難所のみAstra、日常処理は安価なモデルを  
併用する多段設計が合理的



## 主要モデルとのベンチマークおよび価格比較

| 比較項目        | GPT-6 Astra | Claude Fable 5.1 | Gemini 3.8 Flash |
|-------------|-------------|------------------|------------------|
| 総合知能(Index) | 61.2        | 65.7             | 58.7             |
| コーディング能力    | 57.9        | 55.8             | 19.1             |
| 1M入力価格      | \$10        | \$10             | \$0.75           |