

# 中国AIの躍進：Kimi K3の実力と導入への指針



## 競合比較とパフォーマンス：特定領域で「世界首位」



Kimi K3

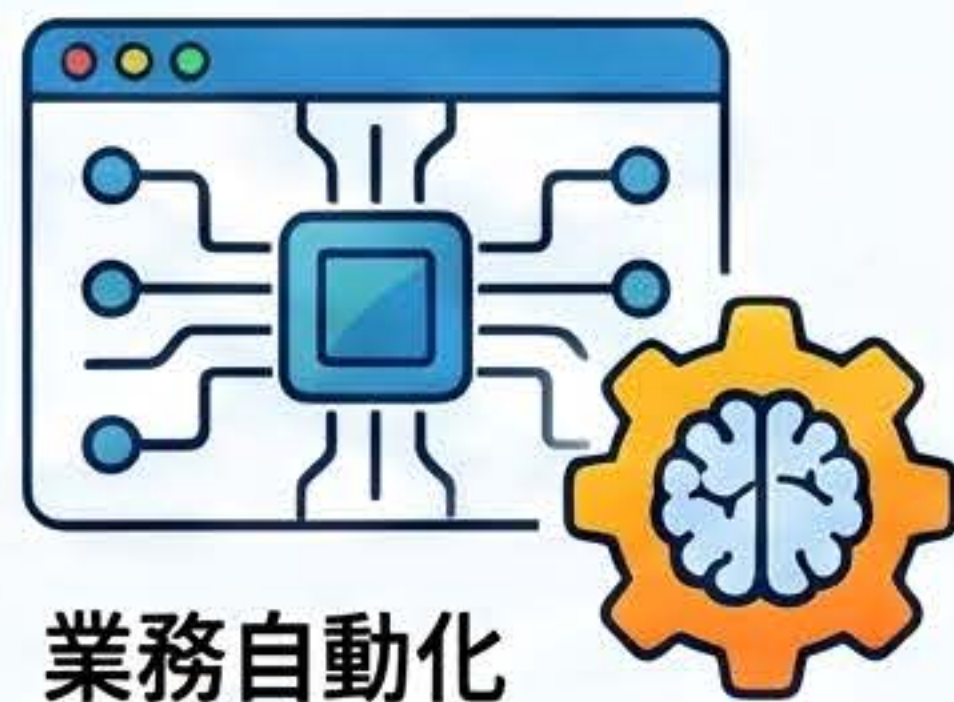
### 総合知能指数で世界第3位

Artificial Analysis Indexで57点を記録し、GPT-5.5 Solを上回るフロンティア層に位置。



### Web開発・コーディングで暫定首位

ArenaのWebDevスコア等、エージェント能力を要する実作業で米国の競合モデルを凌駕。



業務自動化

### 高度な知識推論 (HLE-Full)

推論と知識のトレードオフ。業務自動化に強い一方、高度な知識推論 (HLE-Full) ではFable 5 に大きく差をつけられている。

## 主要なベンチマークにおけるKimi K3と米国勢の立ち位置

| 評価指標                   | Kimi K3   | Fable 5 (Anthropic) | GPT-5.6 Sol (OpenAI) |
|------------------------|-----------|---------------------|----------------------|
| 総合知能指数 (Index)         | 57 (第3位)  | 60 (首位)             | 59 (第2位)             |
| WebDev Overall (Arena) | 1679 (首位) | 1631                | 1618                 |
| SWE Marathon (Code)    | 42 (首位)   | 35                  | 39                   |

## 実務導入における経済性とリスク



### 単価は安い、タスク単価の差は10%

API単価は低い、出力が冗長なため1タスクあたりの実質コストはGPT-5.6 Solと僅差。



### 7月27日に「重み」を公開予定

調査時点ではオープンウェイト「予定」であり、ライセンスや安全性の詳細は公開後の検証が必要。



### 「人間による検証」を前提とした運用

探索やコード生成には有力だが、機密情報や最終判断には他モデルとの併用や人の介入を推奨。