

防衛分野における最先端AIの活用

Motoki Sato

2025/11/11
防衛装備庁技術シンポジウム2025



01

はじめに

02

Sakana AIのR&D

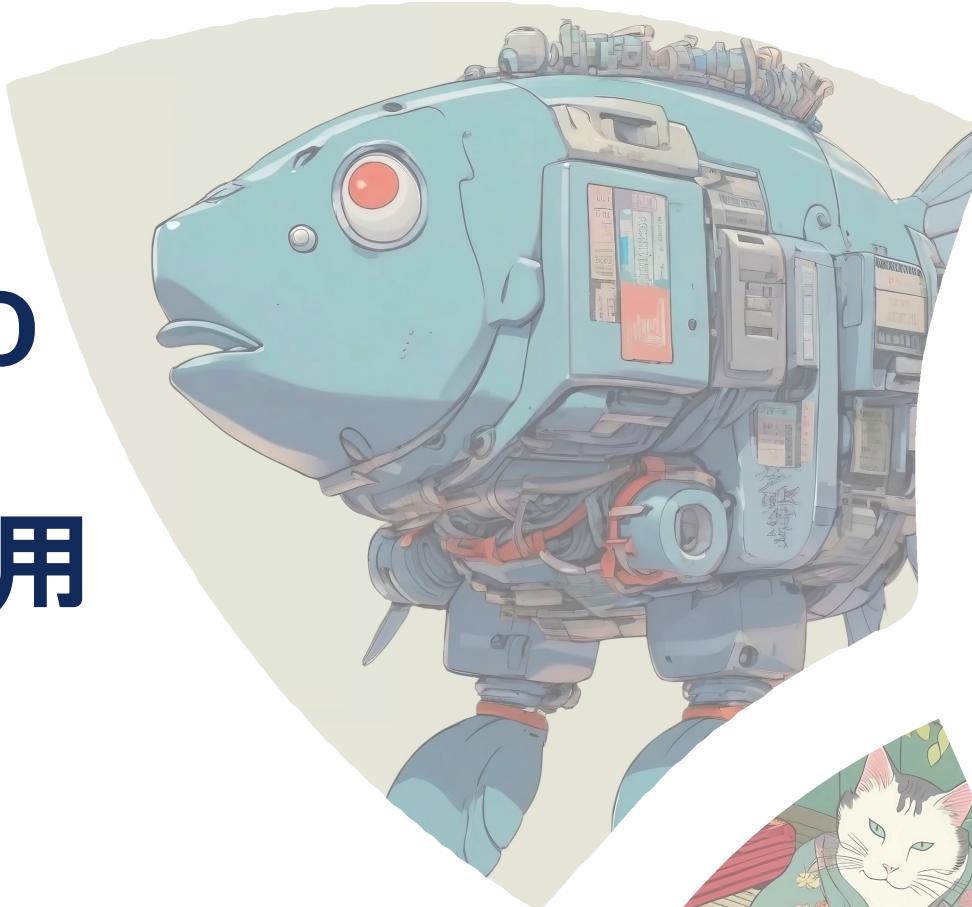
03

防衛分野への応用

04

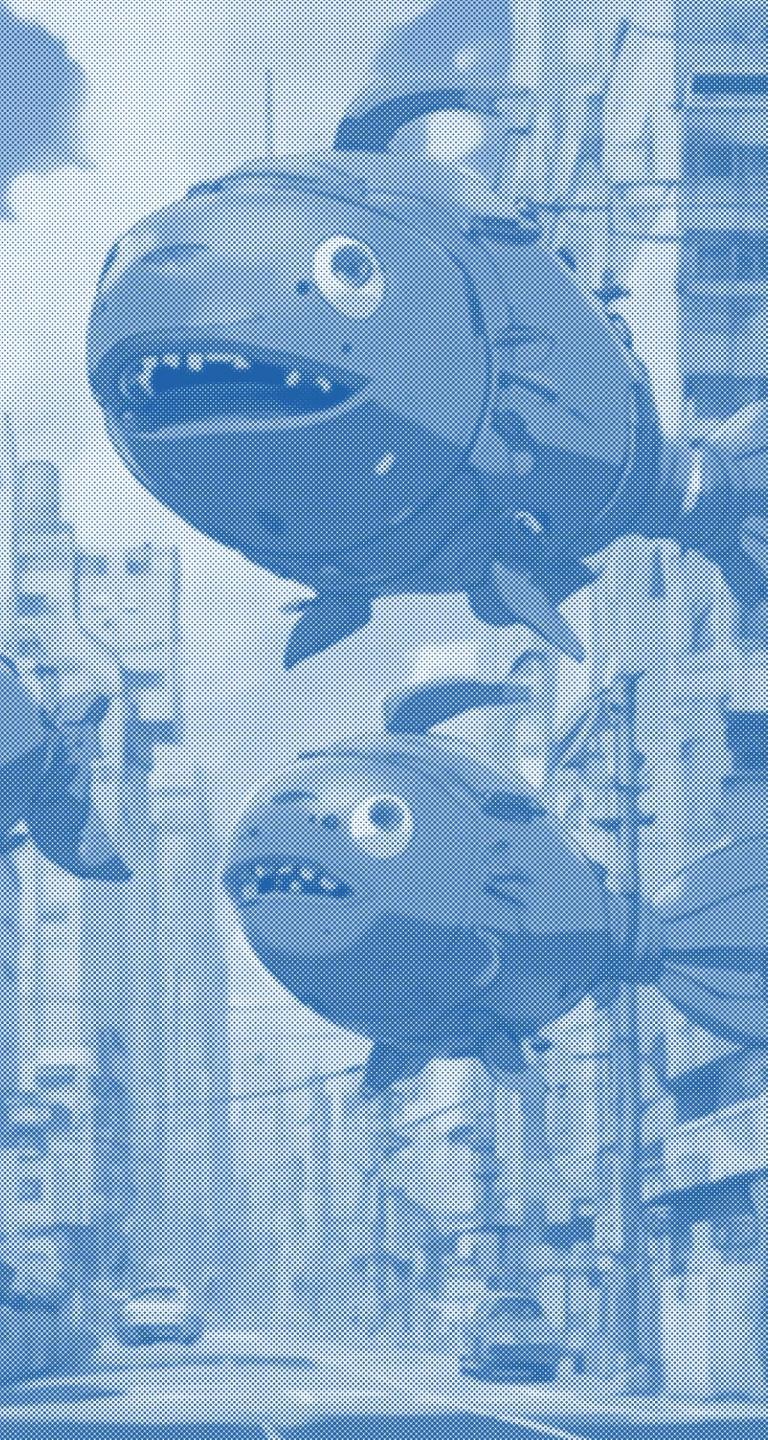
おわりに

AGENDA



01

はじめに Sakana AIのご紹介



会社紹介: Sakana AI



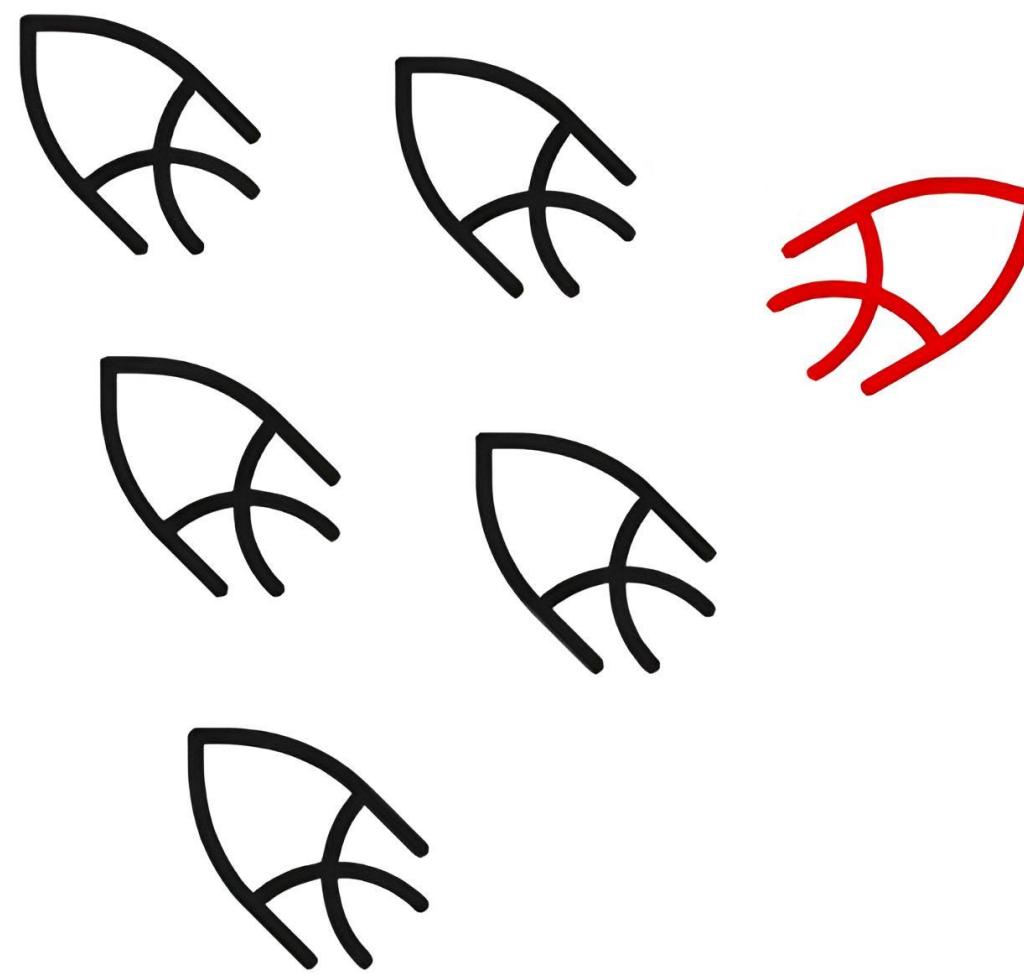
CEO
David Ha
Head of Google Brain Tokyo



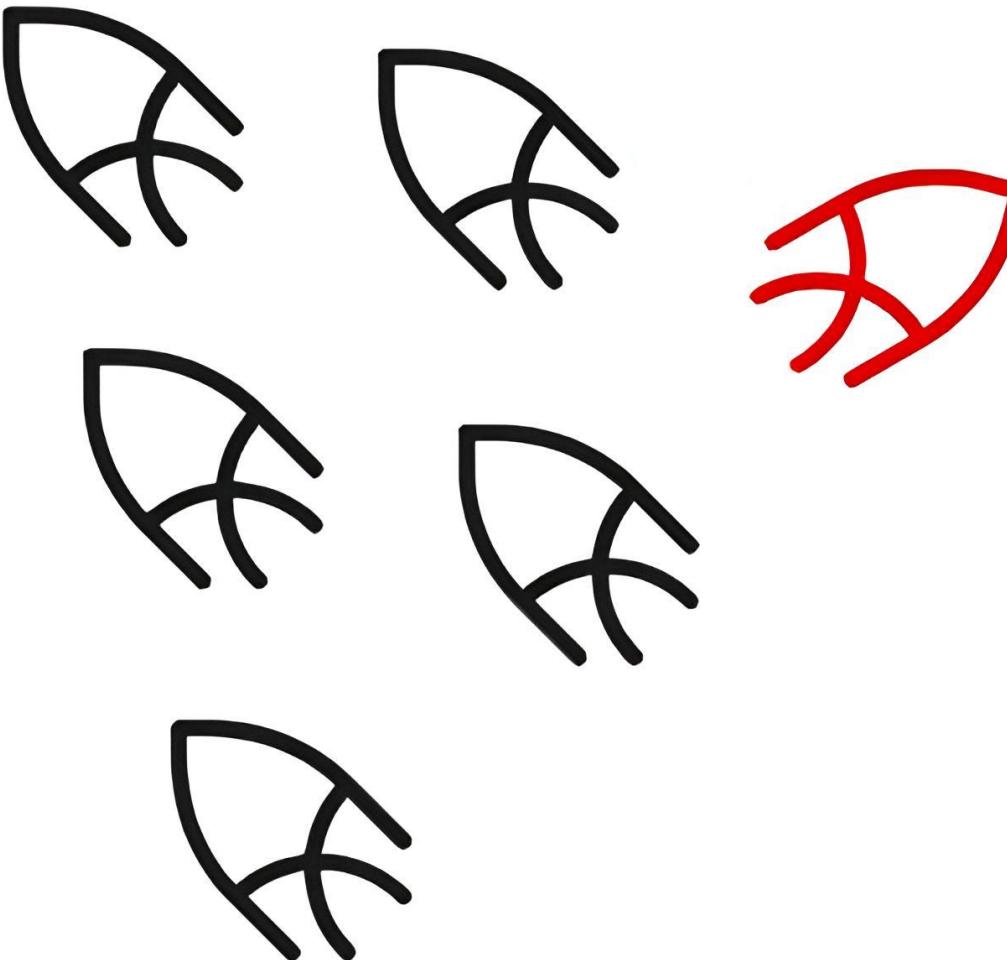
CTO
Llion Jones
Co-invented
“Transformer”



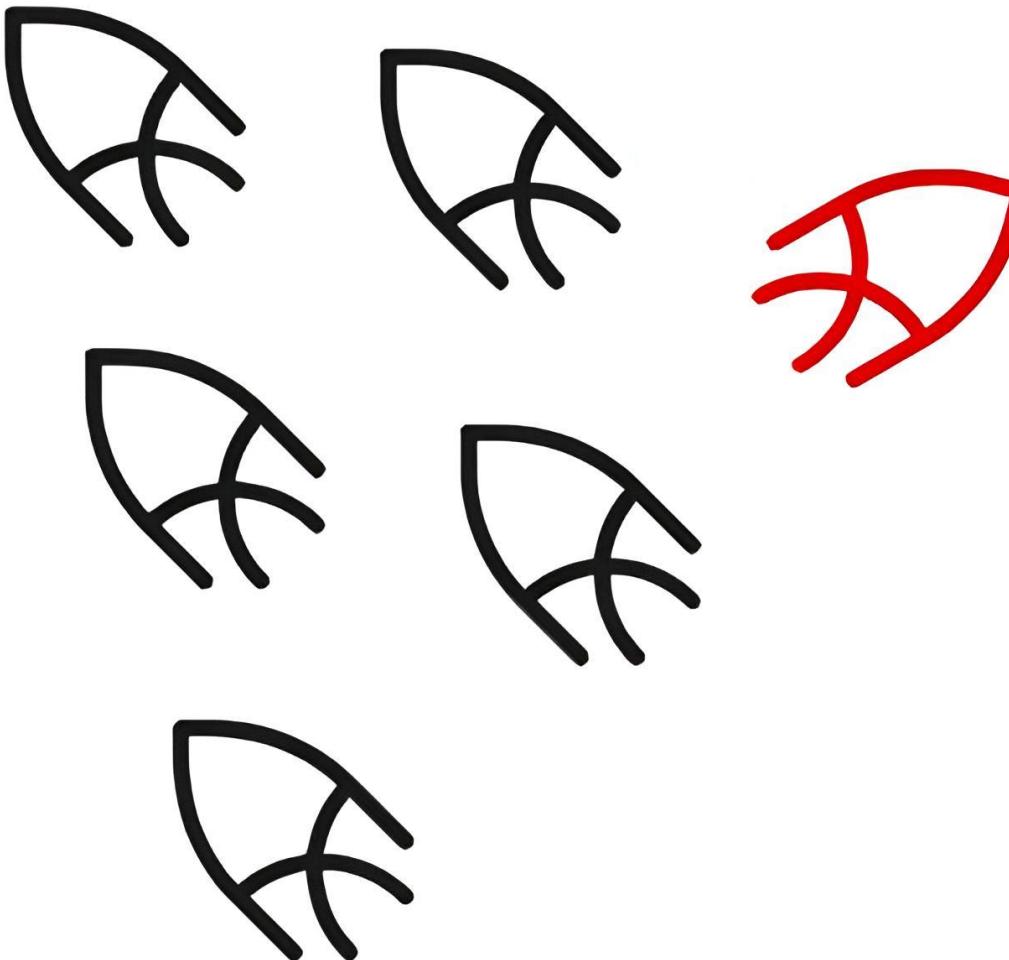
COO
Ren Ito
From diplomacy to tech:
CEO of Mercari Europe.



モデルを大きく

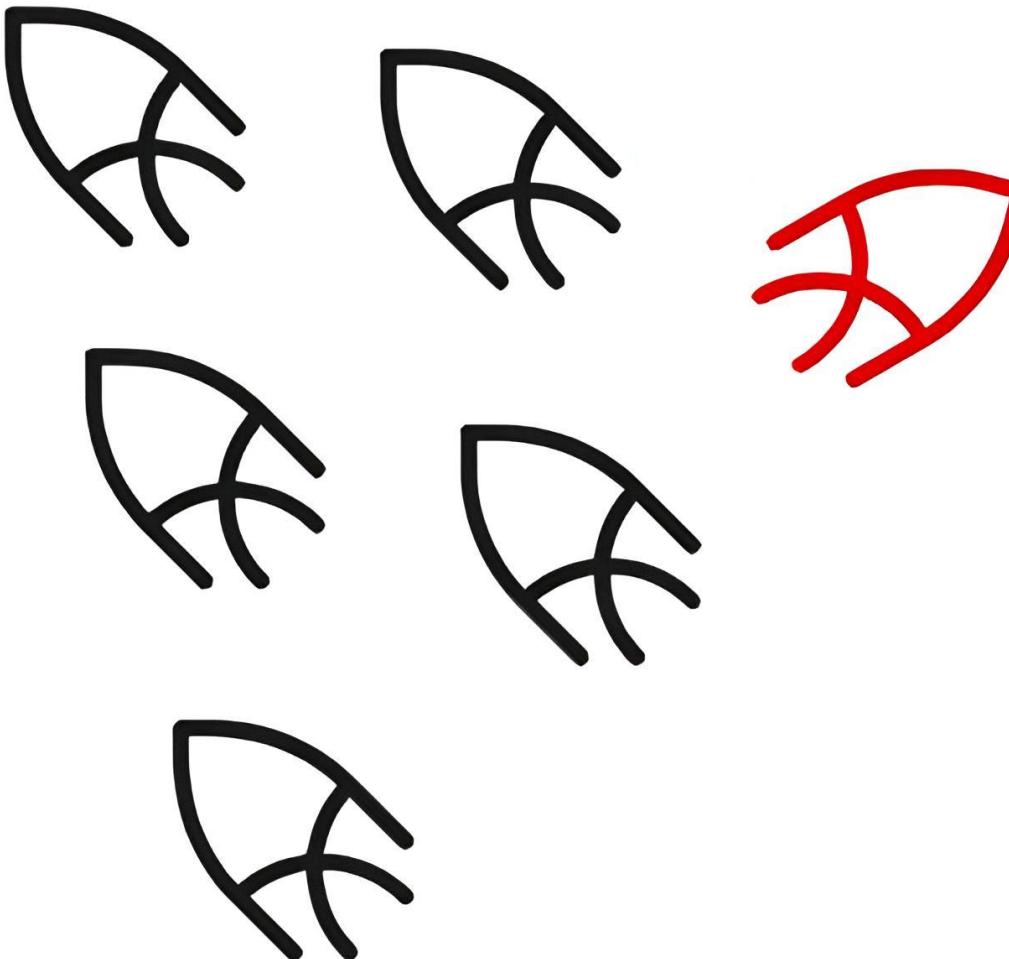


モデルを大きく



データを大きく

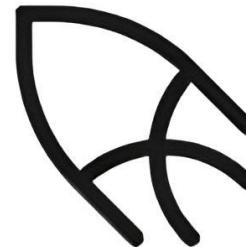
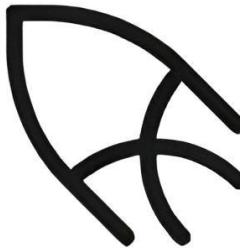
モデルを大きく



データを大きく

学習を長く

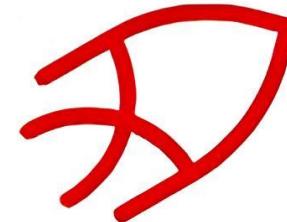
モデルを大きく



データを大きく



学習を長く



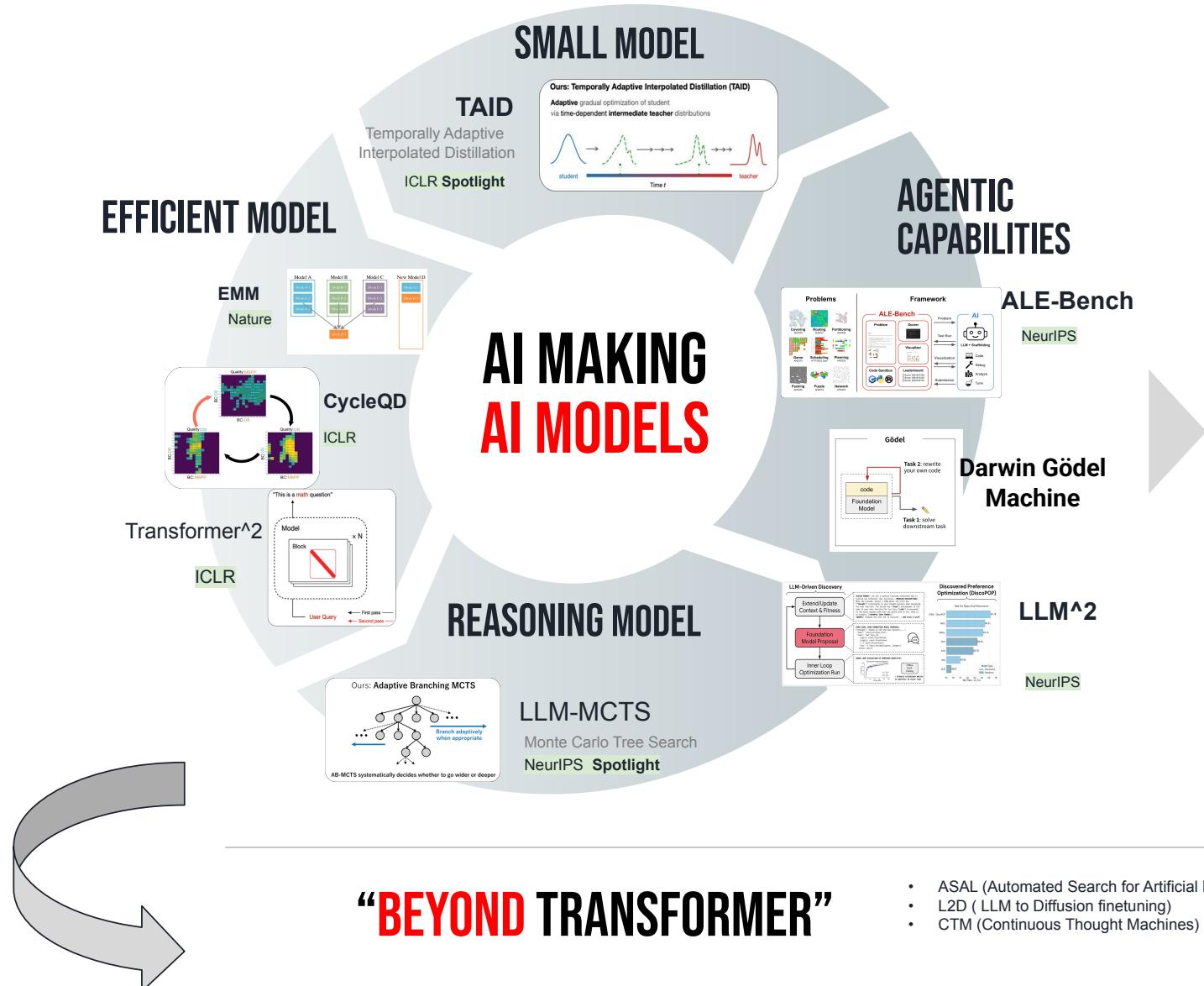
他にもやるべきこと
があるはず

会社紹介: Sakana AI

“The core research focus of Sakana AI is in applying nature-inspired ideas, such as evolution and collective intelligence, to improve foundation models’ performance”



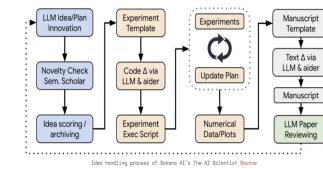
Sakana AIのR&D



本日扱う技術要素

WORKFLOW AUTOMATION

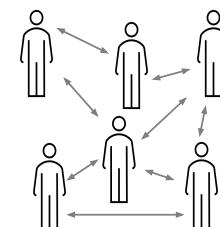
AI Scientist



SMALL MODEL

TAID

エッジで動作するSLMの作成技術



SIMULATION

Agent Based Modeling (ABM)



02

AI Scientist 基盤モデルの研究開発



Sakana AIが開発した自立型エージェント：AI Scientist

2024-9-4

The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery

Chris Lu^{1,2,*}, Cong Lu^{3,4,*}, Robert Tjarko Lange^{1,*}, Jakob Foerster^{2,†}, Jeff Clune^{3,4,5,†} and David Ha^{1,†}

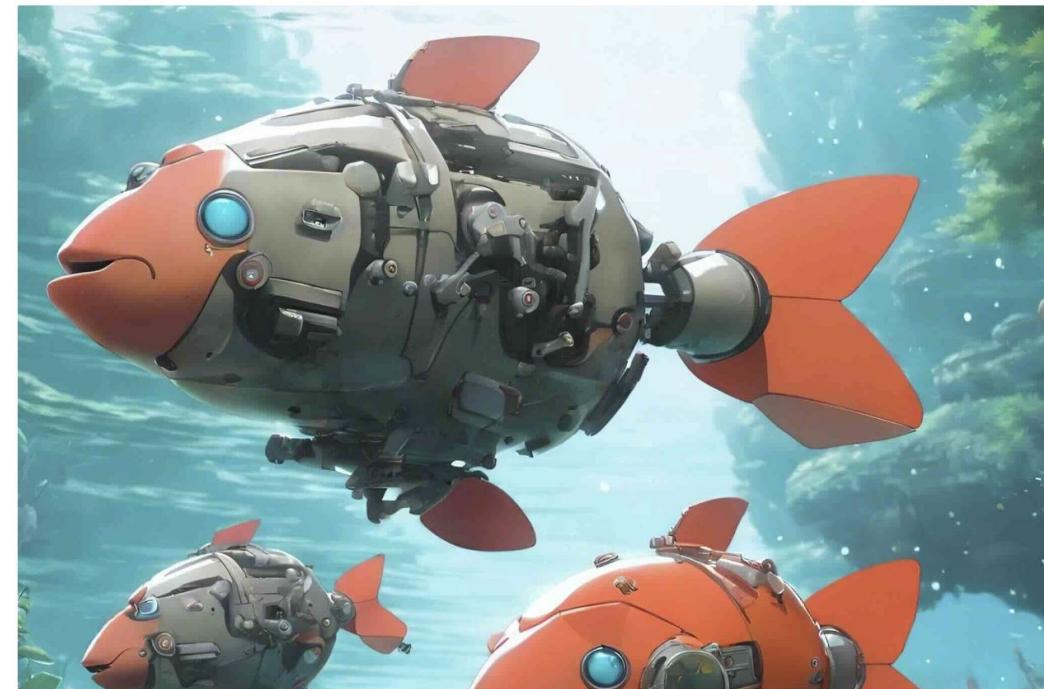
*Equal Contribution, ¹Sakana AI, ²FLAIR, University of Oxford, ³University of British Columbia, ⁴Vector Institute, ⁵Canada CIFAR AI Chair, [†]Equal Advising

One of the grand challenges of artificial general intelligence is developing agents capable of conducting scientific research and discovering new knowledge. While frontier models have already been used as aides to human scientists, e.g. for brainstorming ideas, writing code, or prediction tasks, they still conduct only a small part of the scientific process. This paper presents the first comprehensive framework for fully *automatic scientific discovery*, enabling frontier large language models (LLMs) to perform research independently and communicate their findings. We introduce THE AI SCIENTIST, which generates novel research ideas, writes code, executes experiments, visualizes results, describes its findings by writing a full scientific paper, and then runs a simulated review process for evaluation. In principle, this process can be repeated to iteratively develop ideas in an open-ended fashion and add them to a growing archive of knowledge, acting like the human scientific community. We demonstrate the versatility of this approach by applying it to three distinct subfields of machine learning: diffusion modeling, transformer-based language modeling, and learning dynamics. Each idea is implemented and developed into a full paper at a meager cost of less than \$15 per paper, illustrating the potential for our framework to democratize research and significantly accelerate scientific progress. To evaluate the generated papers, we design and validate an automated reviewer, which we show achieves near-human performance in evaluating paper scores. THE AI SCIENTIST can produce papers that exceed the acceptance threshold at a top machine learning conference as judged by our automated reviewer. This approach signifies the beginning of a new era in scientific discovery in machine learning: bringing

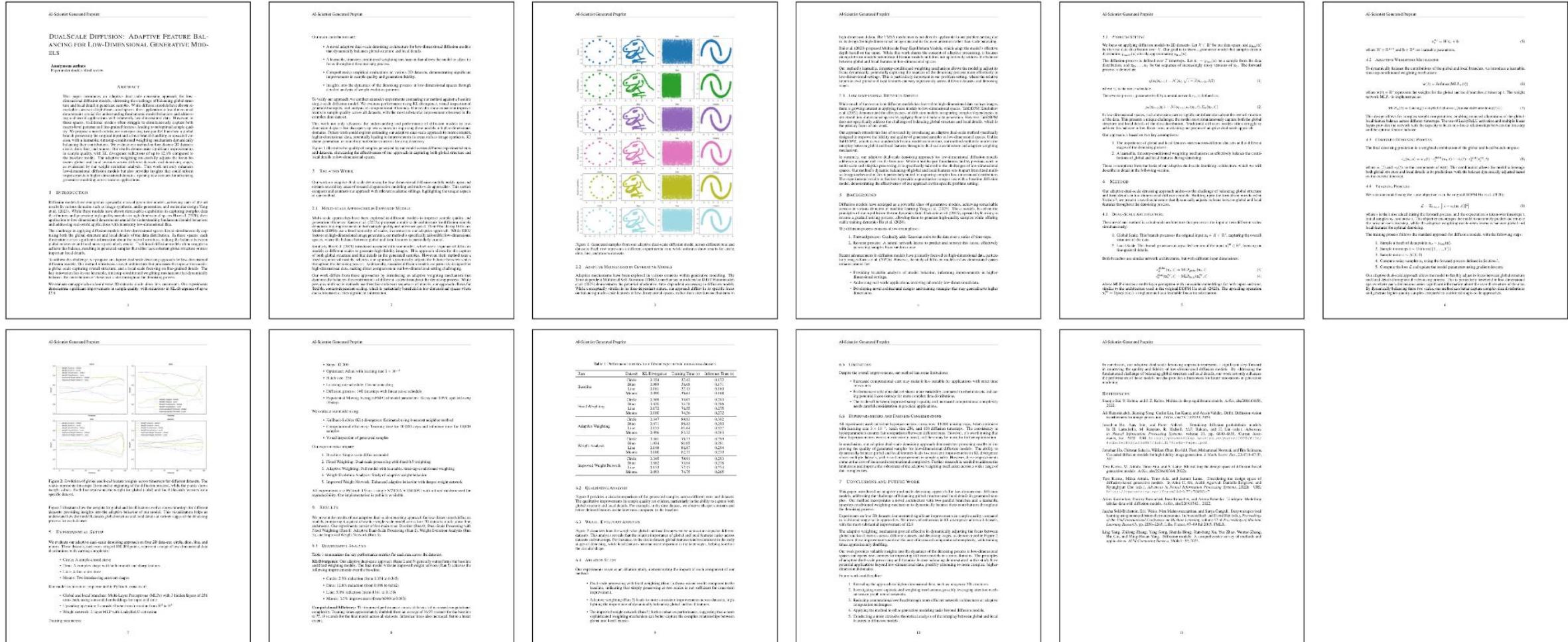
sakana.ai

The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery

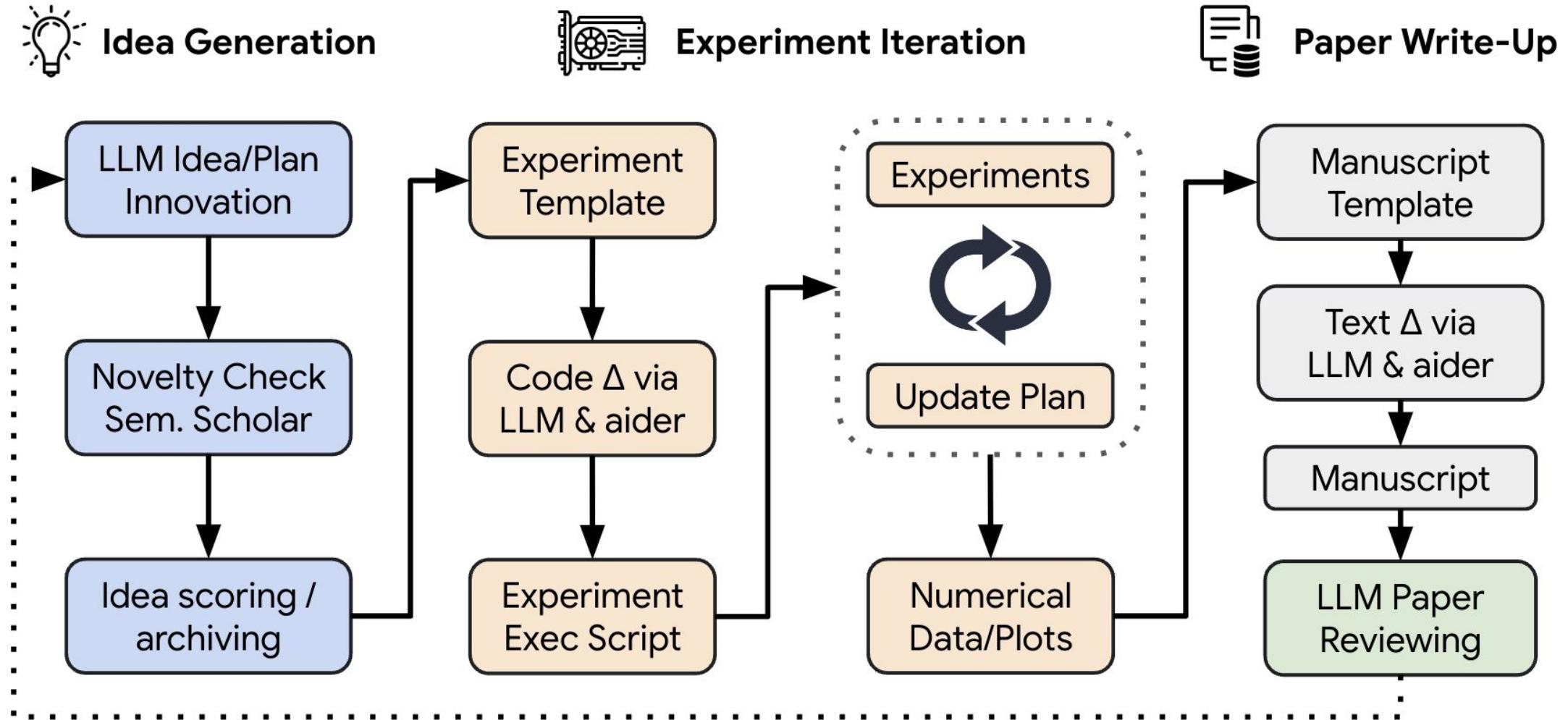
August 13, 2024



Sakana AIが開発した自立型エージェント：AI Scientist

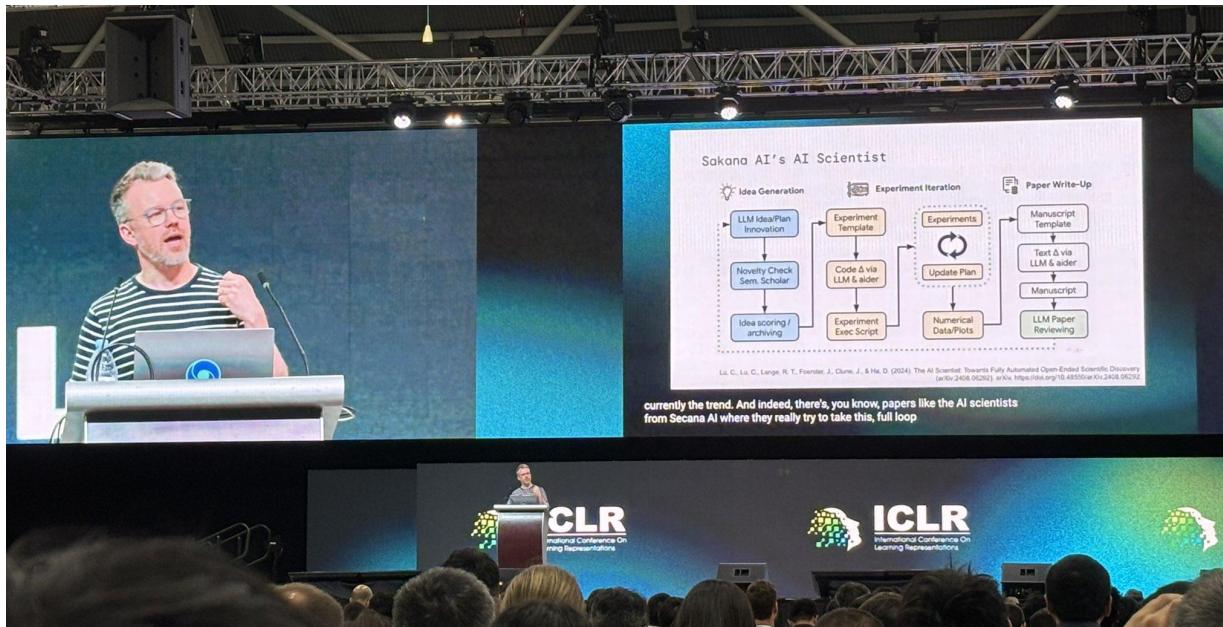


Sakana AIが開発した自立型エージェント : AI Scientist



Sakana AIが開発した自立型エージェント：AI Scientist

- トップレベル学会のワークショップで **査読プロセスを通過**
 - 完全AI生成論文が人間の査読プロセスを通過した**世界初の事例** (2025年4月)
- 学会基調講演や Natureなどのジャーナルにて大々的に言及される

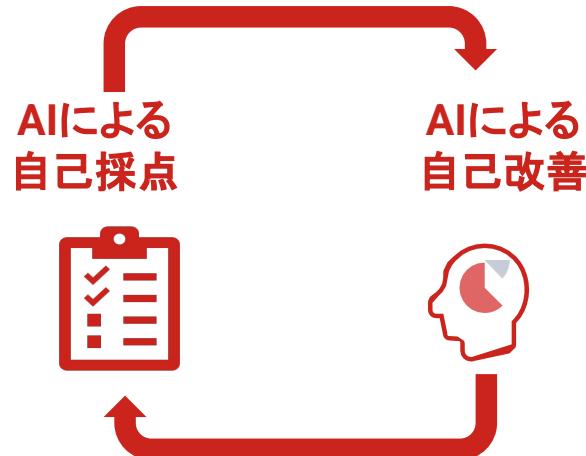


A screenshot of a news article from the website "nature". The title is "Researchers built an 'AI Scientist' – what can it do?". The article discusses how the large language model can perform tasks from reading literature to writing and reviewing its own papers, but notes limited applicability. It includes a small illustration of a blue robot scientist holding a test tube.

エッセンス (1/2)

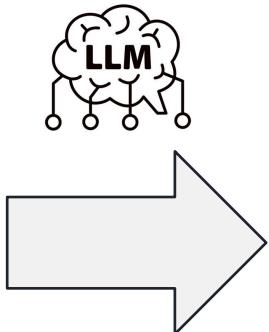
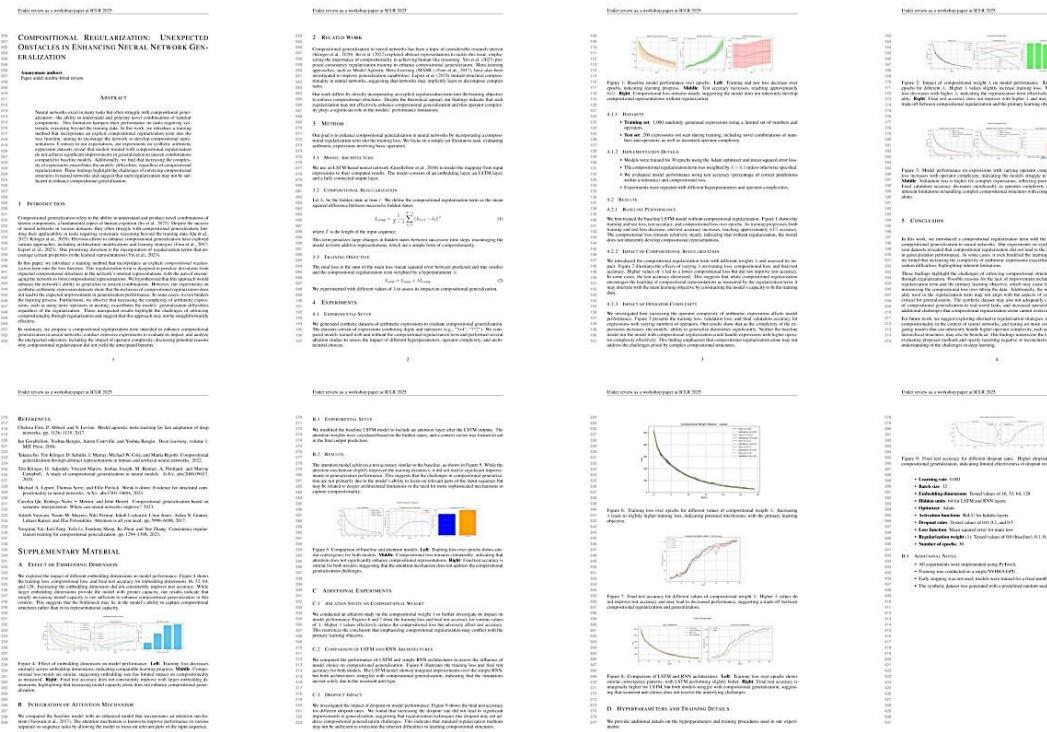
- 実験基盤を内包した **進化計算** の活用

- 試行錯誤しては微修正...という、自然界に着想を得た自己改善プロセス
- 新たな知見を探索するための技術である「**オープンエンド探索**」に該当



エッセンス (2/2)

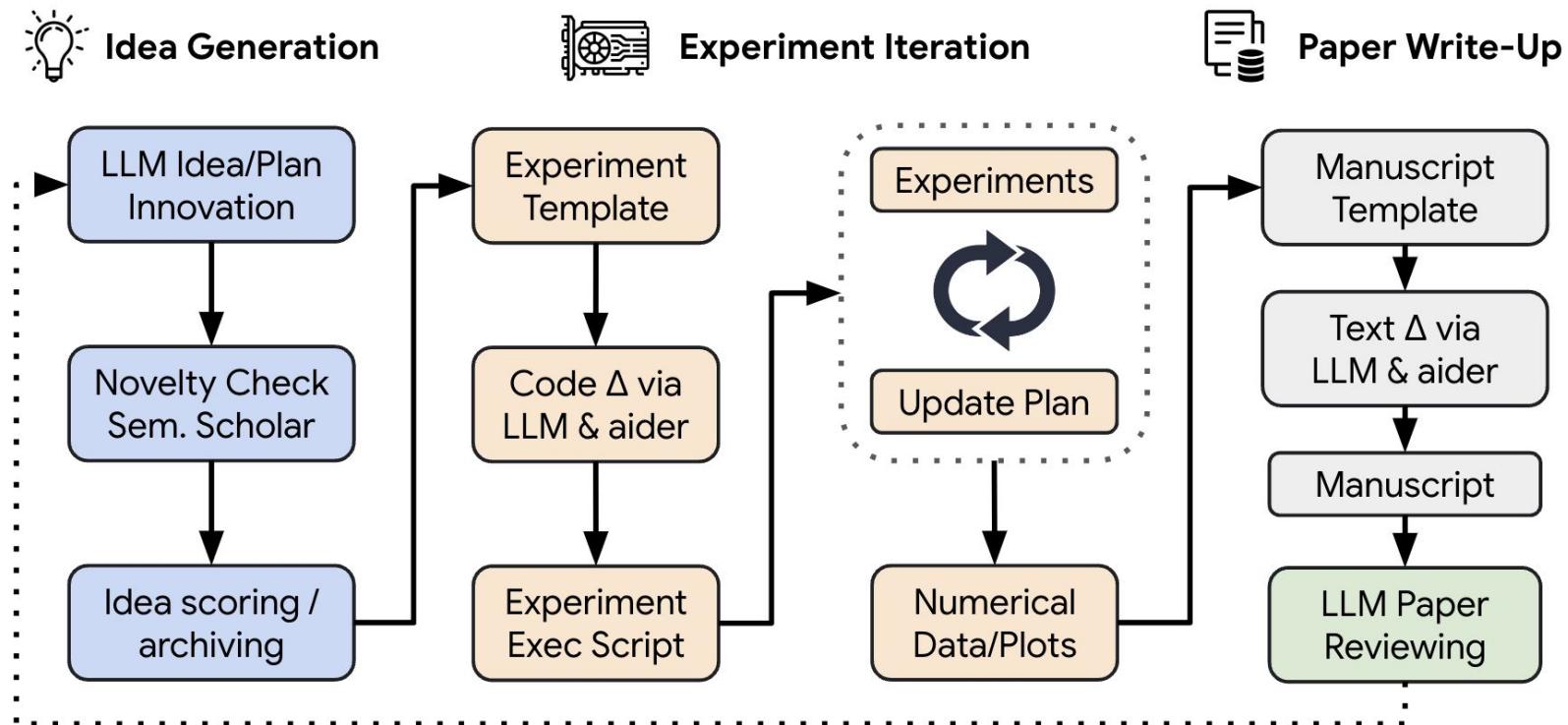
- 定量評価しにくいものの 定量化
 - 「LLM as a Judge」と呼ばれる技術の高度化



10段階における 評価は6

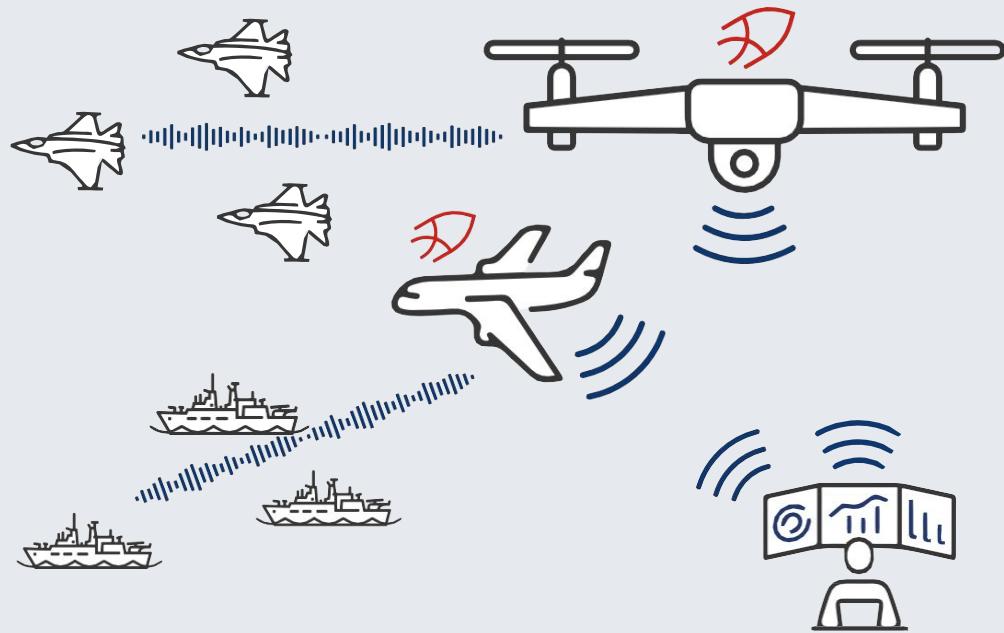
幅広い活用可能性

- 汎用的なフレームワーク であり、多様な活用チャンスがある
 - 「着想→試行→まとめ→次の手...」というのは、至る所に現れる枠組み



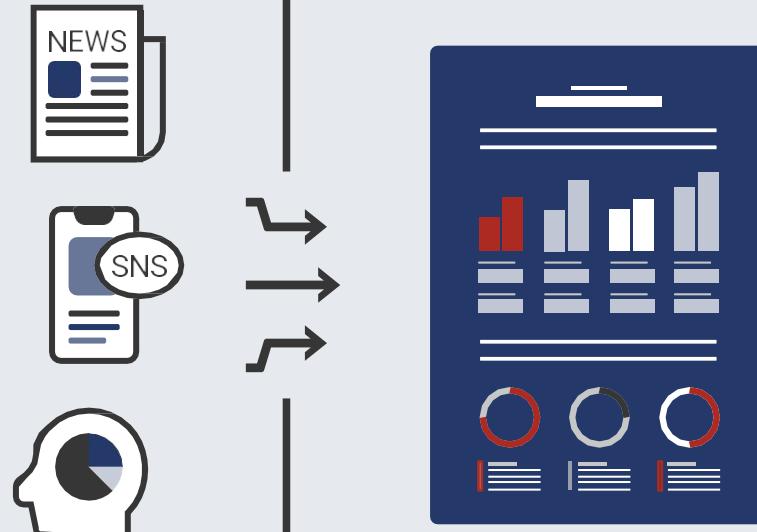
防衛分野における取り組み

無人機の自律的制御



無人アセット (UAV, USV, UUV etc)をAIを用いてエッジ側で自律的に制御することに取り組む。

インテリジェンス分析



偽情報対策や認知戦を含むオープンソース分析による国家インテリジェンス強化に取り組む。

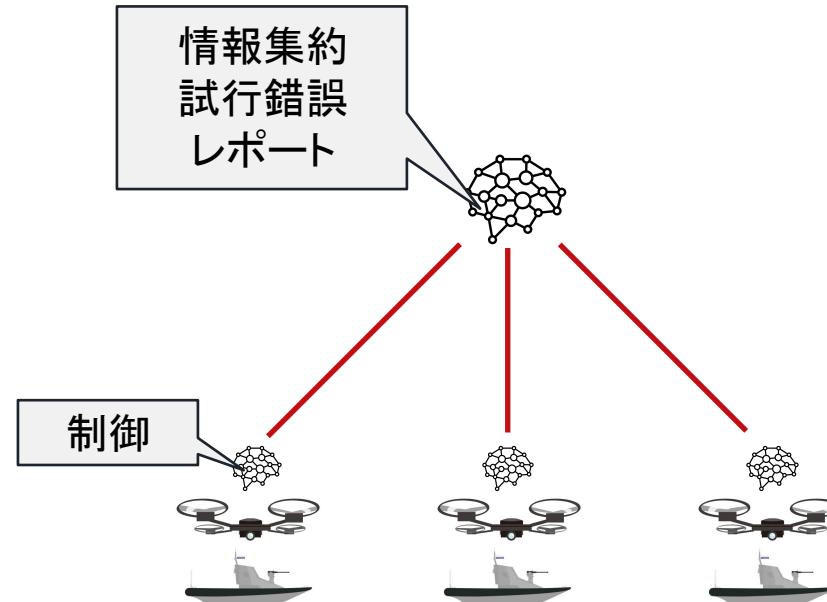
03

防衛分野における取り組み 基盤モデルと制御システムについて



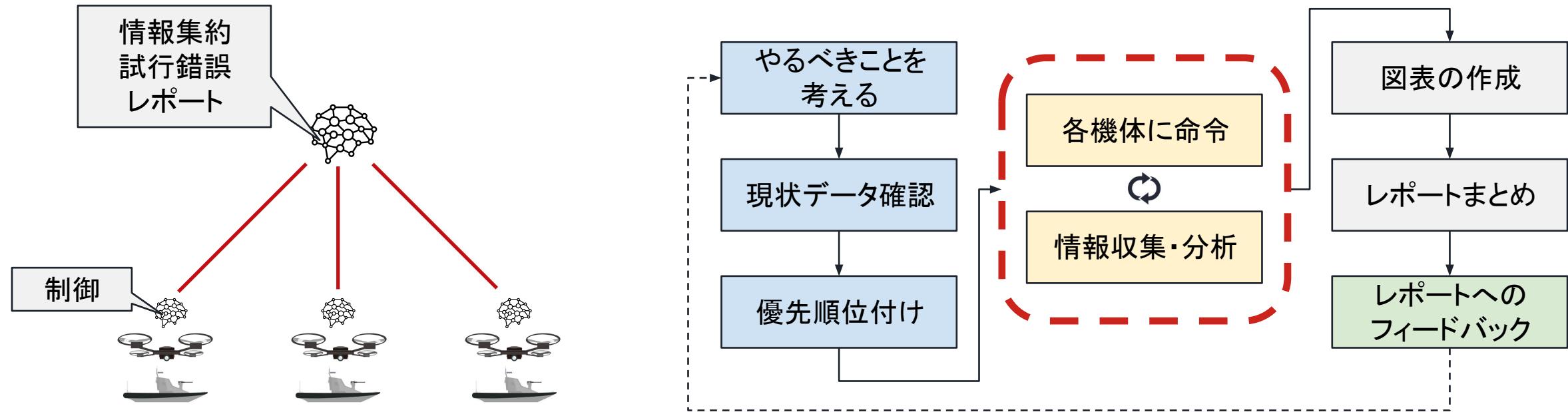
制御システムと AI Scientistフレームワーク

- 無人機を用いた監視・点検・調査の精緻化・緻密化
 - 制御まで含めた仕組みへの拡張



制御システムとAI Scientistフレームワーク

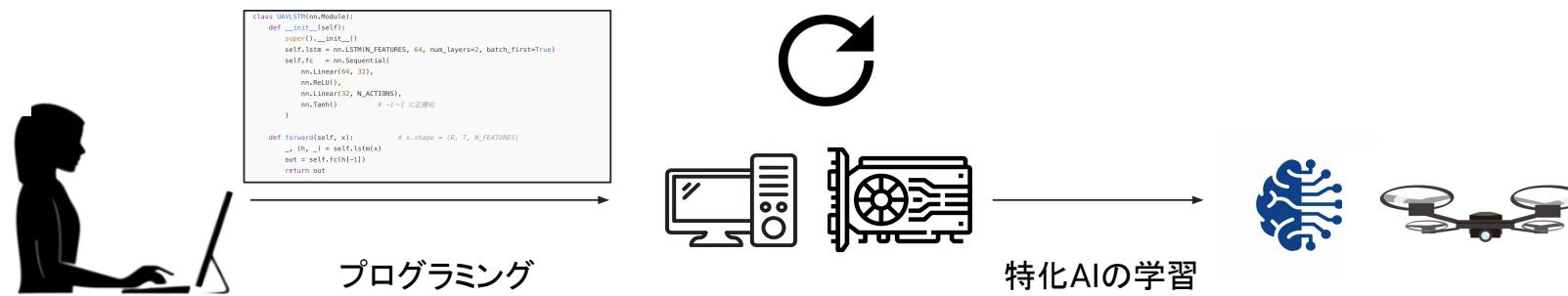
- 無人機を用いた監視・点検・調査の精緻化・緻密化
 - 制御まで含めた仕組みへの拡張



無人機の自律的制御における課題

従来はタスク毎・機種ごとに特化 AIを学習する必要 があり量産に耐えられない

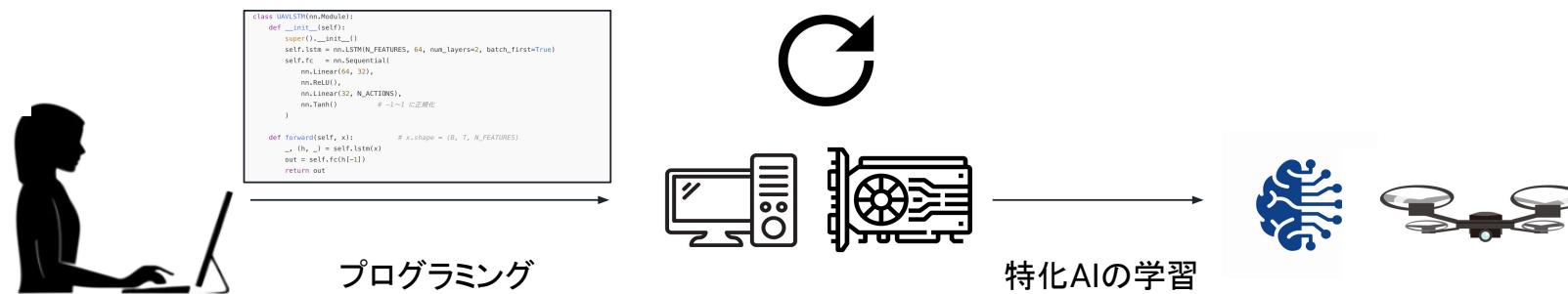
- 学習された状況には柔軟に動作するが、高度な開発技術や計算資源が必要
- 環境などが変化すると、また作り直す必要がある



無人機の自律的制御における課題

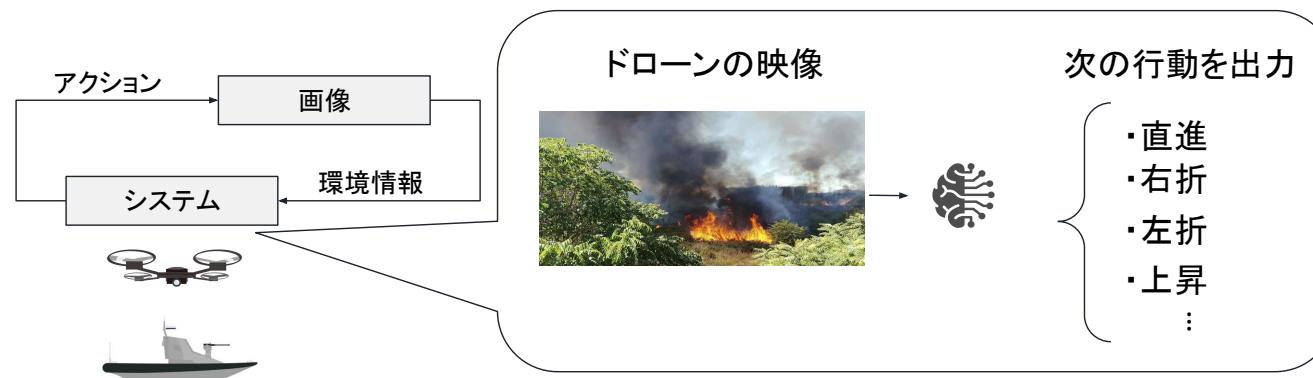
従来はタスク毎・機種ごとに特化 AIを学習する必要 があり量産に耐えられない

- 学習された状況には柔軟に動作するが、高度な開発技術や計算資源が必要
- 環境などが変化すると、また作り直す必要がある

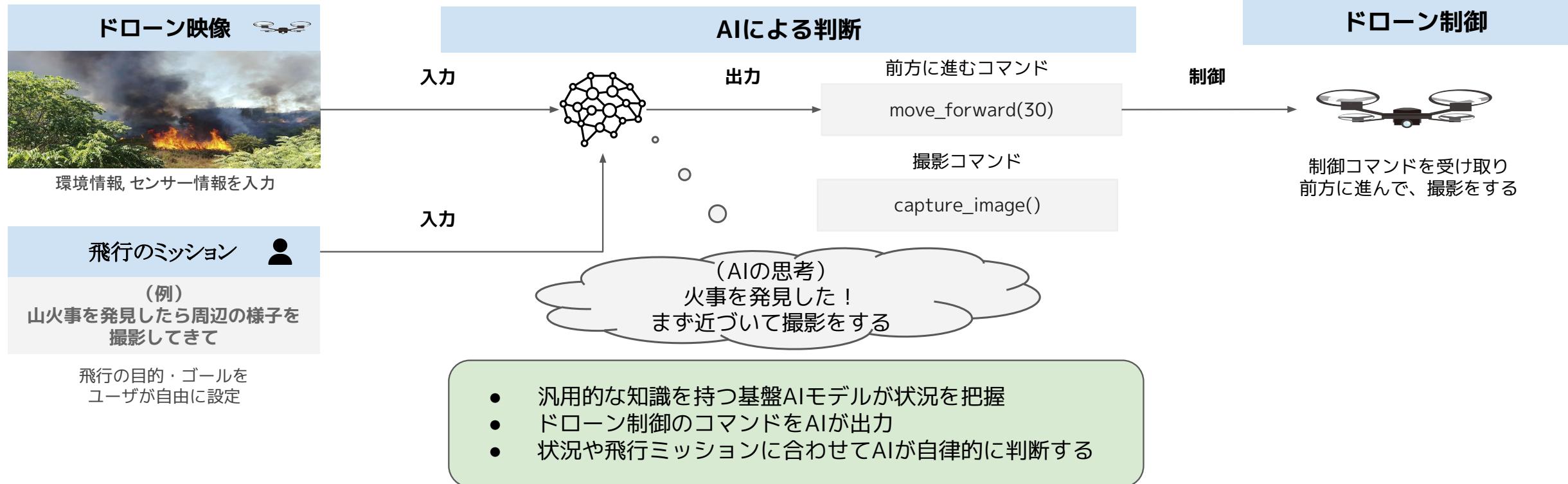


基盤モデルの強みを活かした、無人機の制御機能開発の効率化 に取り組む

- カメラ映像を入力として、ドローン制御コマンドを出力する機能を開発する
- 環境の変化へ柔軟に対応が可能で機種依存しないアプローチ



提案ソリューション



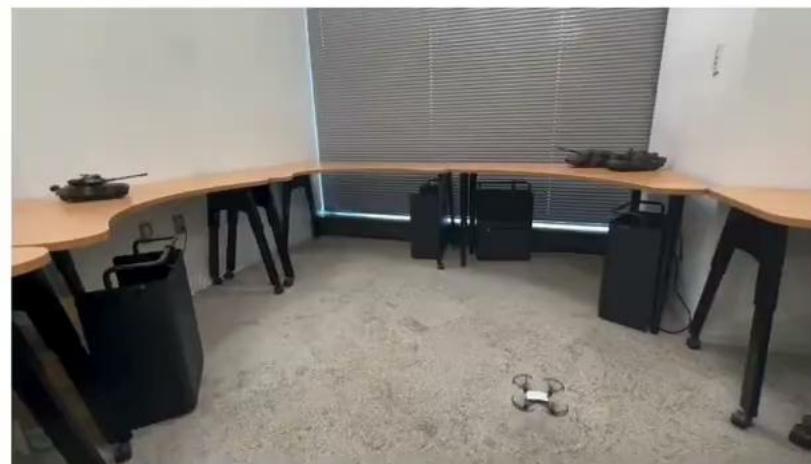
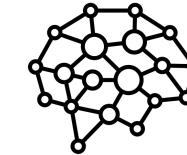
- **Sakana AIは特化モデルでなく、基盤モデルを活用した制御 を実施する**
 - 基盤モデルが持つ汎用的な知識を活用し、状況に応じてAIが自律的に考え実行する
 - **任意のタスクに任意の無人機で 取り組める**
 - 特定の無人機に特化せず、既存の無人機も活用しながら汎用的な制御を行う

デモ動画

ミッション



戦車を見つけ、なるべく近づき観測してください。
その後、旋回し、他の戦車が無いか探索してください。
その後、探索を終えたら着陸してください。



まだAI分析結果がありません。ドローンに指示して分析を開始してください。

部屋の様子

ドローン映像

AIの判断

無人機 x AIの応用可能性

無人機の自律制御・協調動作に応用が可能

協調動作



協調動作



UAV



自律動作

エッジ上で動作するSLM

AI for Any Mission



UGV



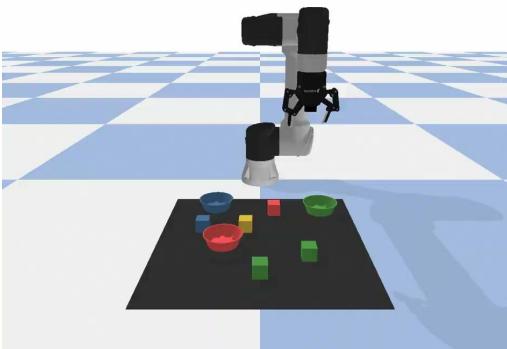
UUV

USV

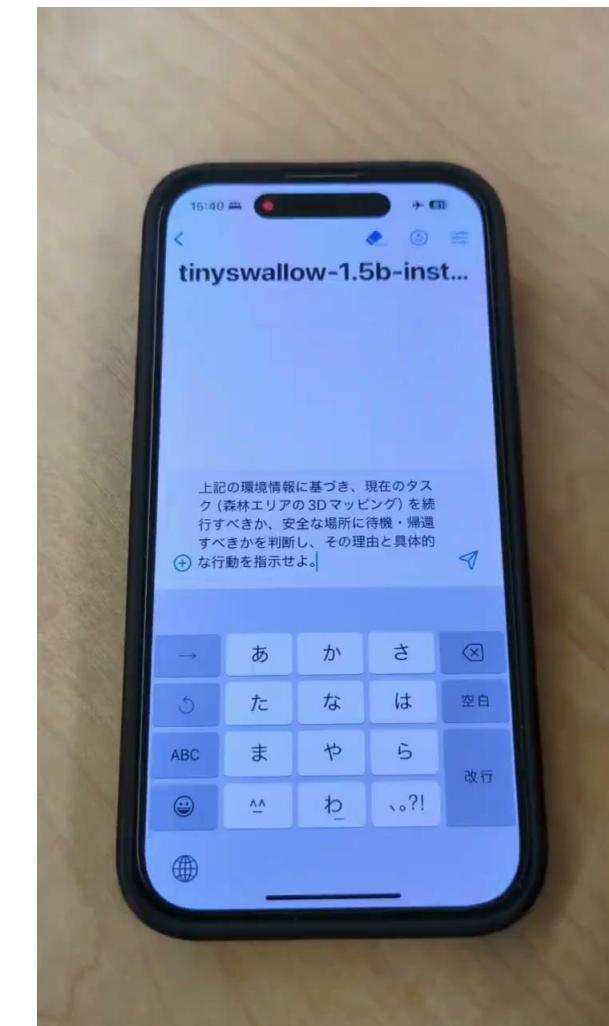
協調動作

Sakana AIの多様な研究開発

- 小型デバイスで動作する**超小型モデル構築技術**
 - 小型日本語最高性能モデルTinySwallow-1.5Bの構築
- 公開モデルと進化計算を用いた**低コスト学習**
 - 進化的モデルマージ
- AIの集団を束ねるためのフレームワーク開発
 - Shachi : <https://github.com/SakanaAI/shachi>
複数機体への応用
- 口ボット制御への応用実験



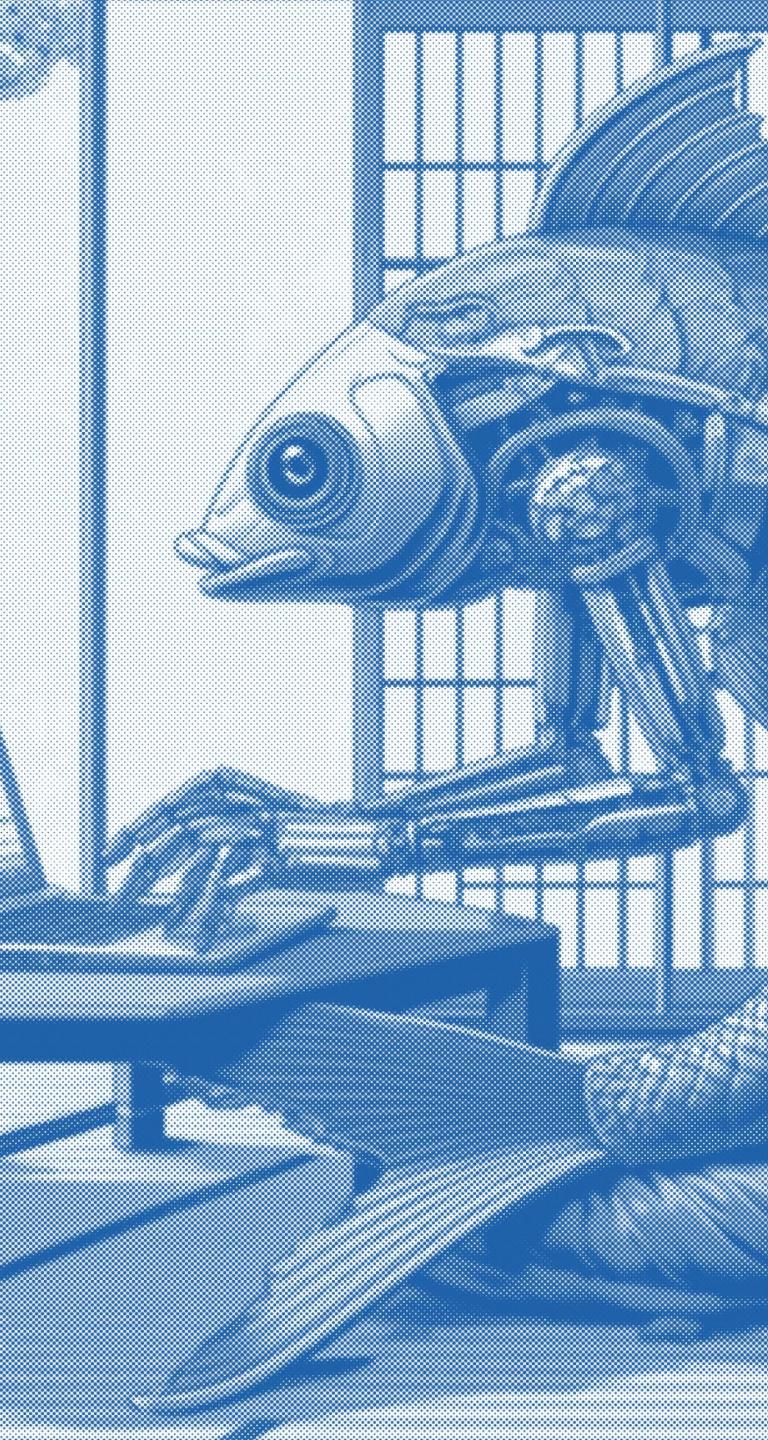
通信がない状況でも動作する SLM



動画: <https://sakana.ai/taid-ja/>

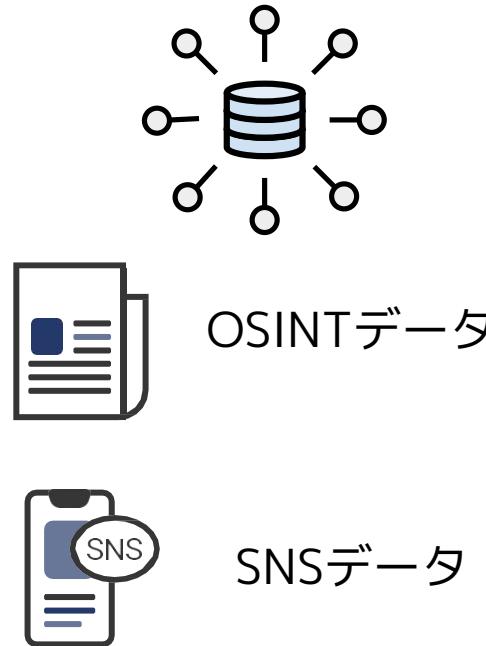
04

防衛分野における取り組み インテリジェンス分析



インテリジェンス分析 新たなる課題への対処

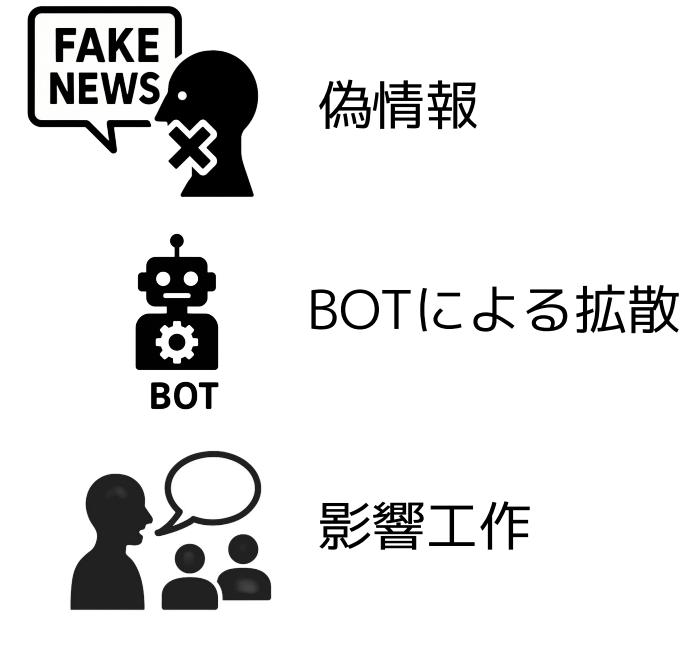
ネット上の 言論空間



影響



課題

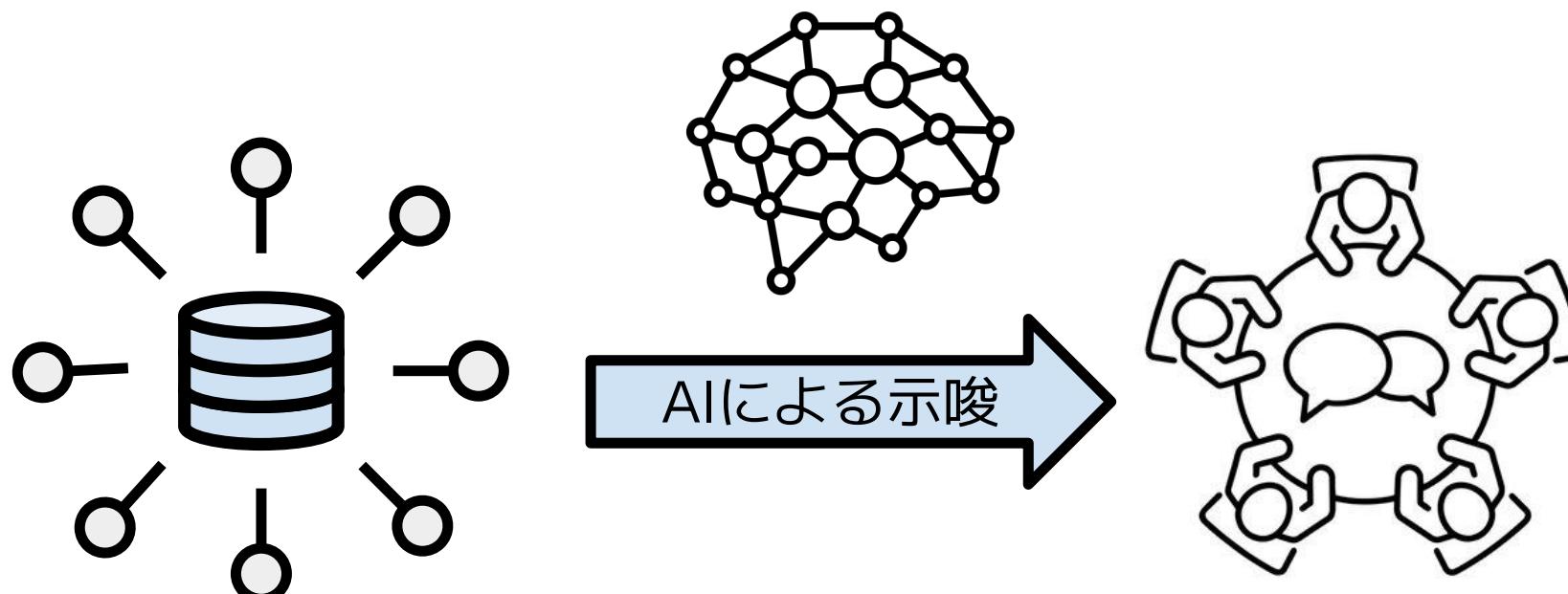


ネット上の言論空間 は 政治・安全保障を分析する上で欠かせない情報ソース

膨大な上、刻一刻と変化するため、伝統的な手法による分析には限界がある

インテリジェンス分析 AIによる強化

AIによる情報分析により、高度な洞察を引き出し意思決定を補助

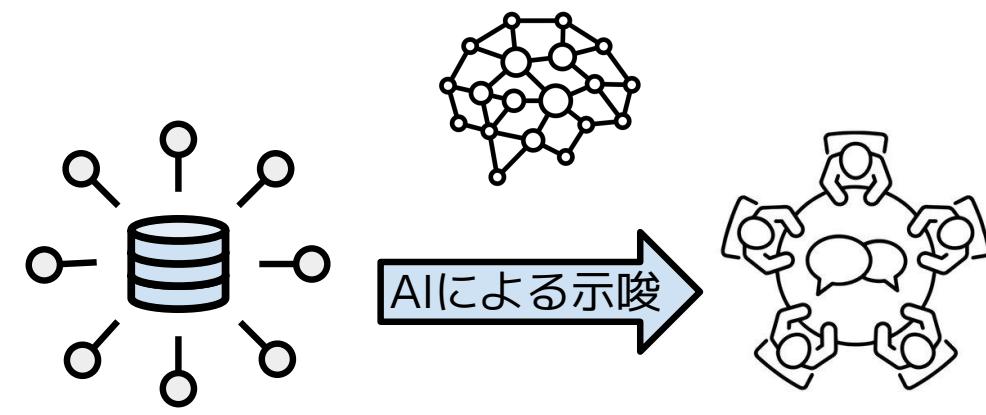


SNSなど、多様な情報を取得

- 意思決定補助
- 脅威への対策
- 現状把握
- ...

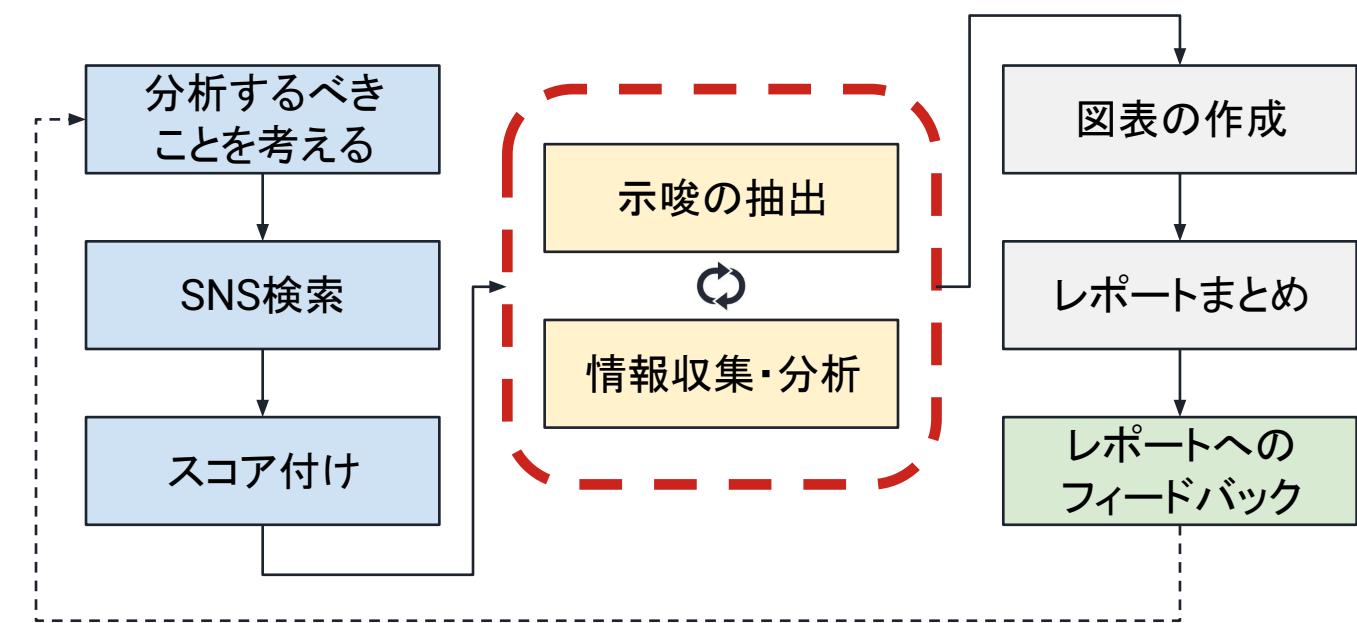
インテリジェンス分析と AI Scientistフレームワーク

AIによる情報分析により、高度な洞察を引き出し意思決定を補助



SNSなど、多様な情報を取得

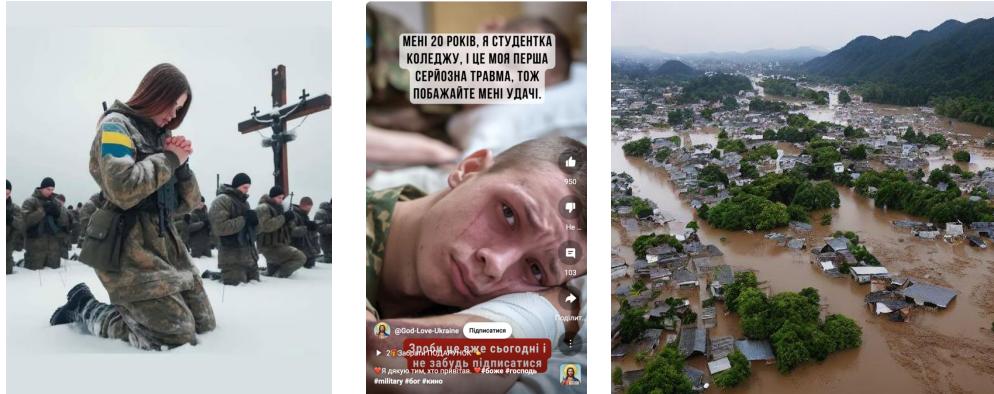
- 意思決定補助
- 脅威への対策
- 現状把握
- ...



AI Scientistのフレームワークが応用可能

偽画像 / 偽情報の検出

AI-Generated Images (生成された画像)



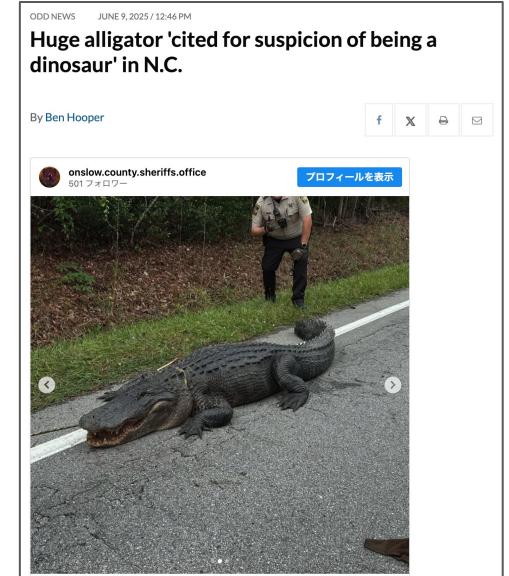
Real Photographs (実際の画像)



SNS上の投稿



AIによる自動検索結果



[AIによる偽情報の自動判定]

Unreliable

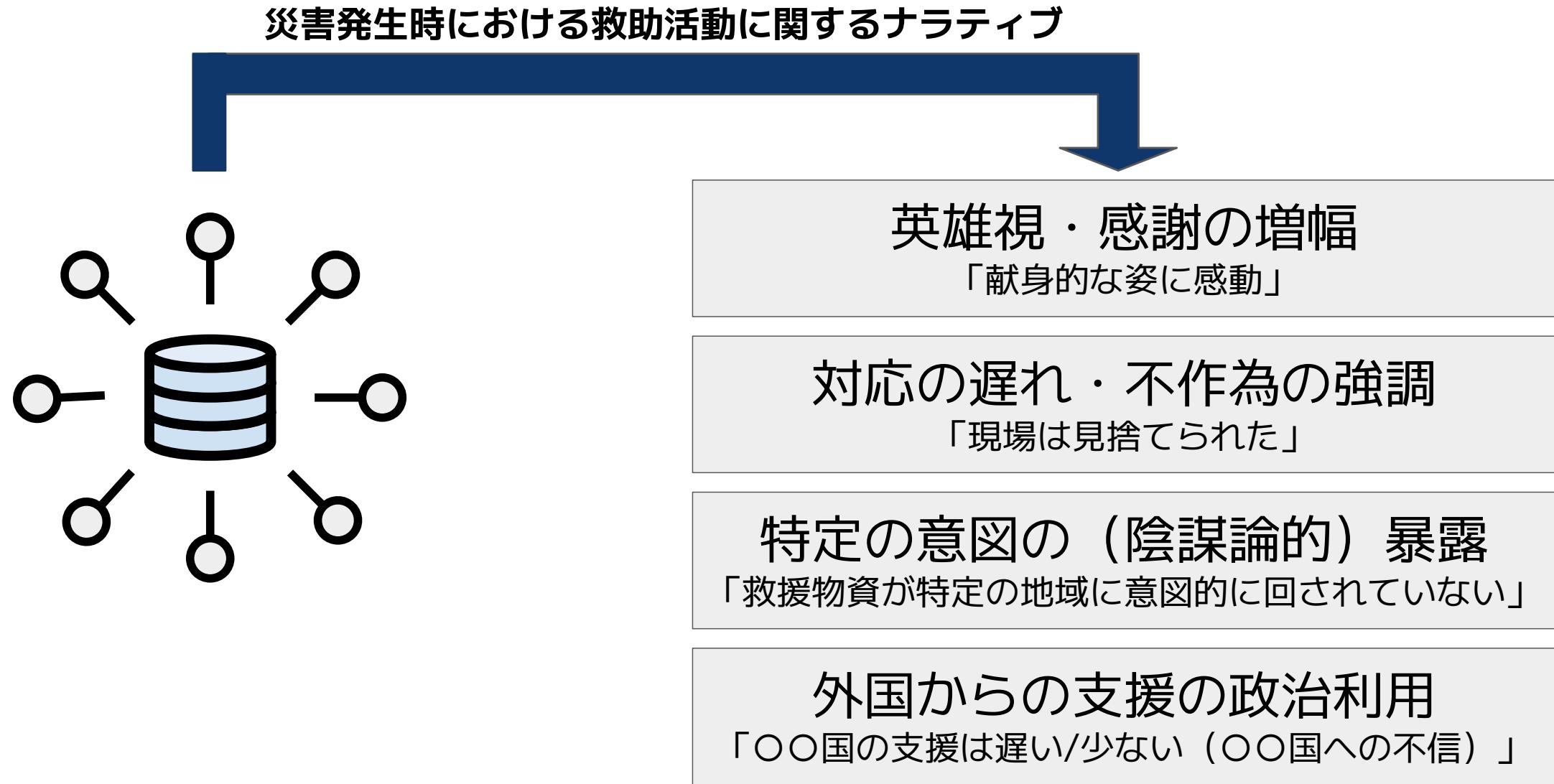
投稿文と画像の説明文を比較した結果、内容は整合していません。

理由:

- ・投稿文: ワニが発見された場所を「都内」と明記しています。
- ・AI探索結果: この画像は「ノースカロライナ州オンスロー郡にある『Old 30 Road』という道路上」で起きた出来事である可能性が高いです。

以上の結果、「都内」と「ノースカロライナ州」は全く異なる場所であり、この点で投稿文と画像は矛盾している可能性が高いです。

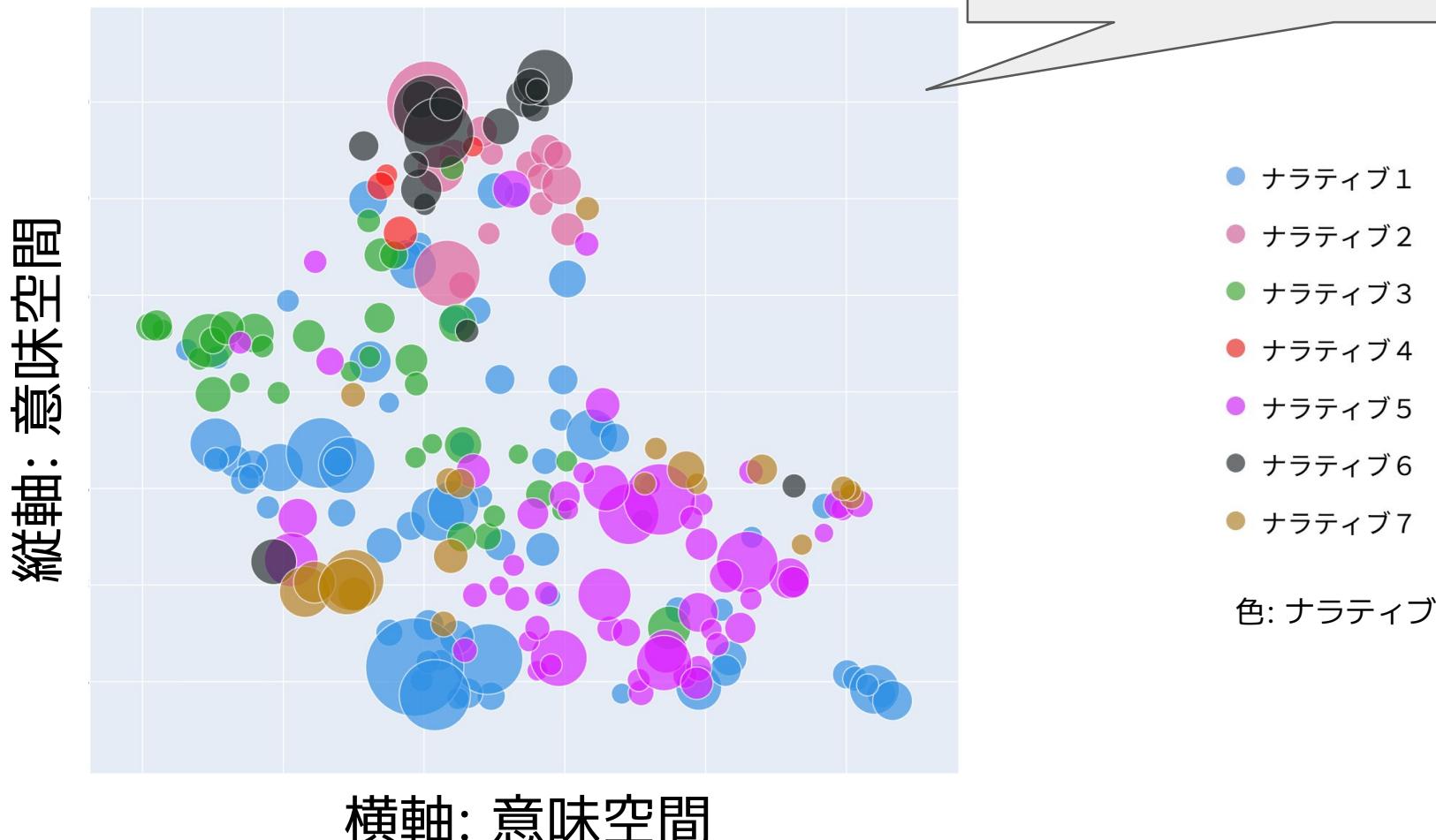
SNSから、特定テーマに関する言及内容（ナラティブ）を抽出



ナラティブ間の距離・勢力の可視化

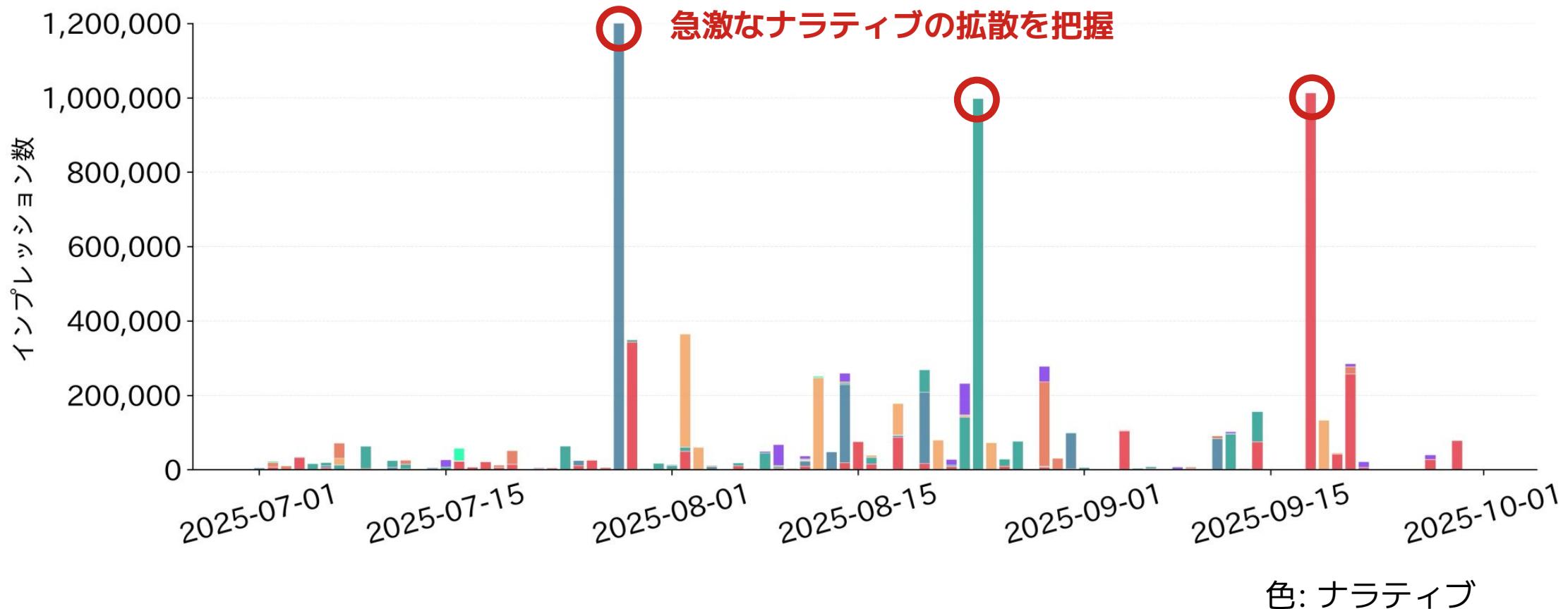
- ナラティブを意味空間によって様々なカテゴリに分解（クラスタ化）
- 各投稿をクラスタによって色分け、拡散力の強さをサイズで可視化
- クラスタ同士の意味空間における近さも可視化
- クラスタの傾向や拡散力を視覚的に把握することが可能に

ナラティブの規模感や、
ナラティブ間の近さを可視化



急激なナラティブの拡散を検知

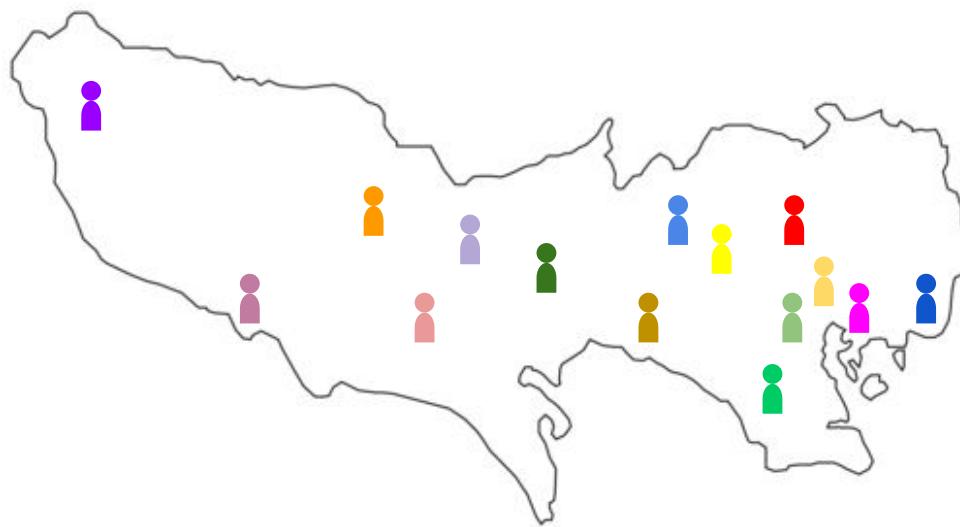
- ナラティブの発生をクラスタごとに可視化
- 特定のクラスタの拡散を迅速に検知、アラートを出すことが可能



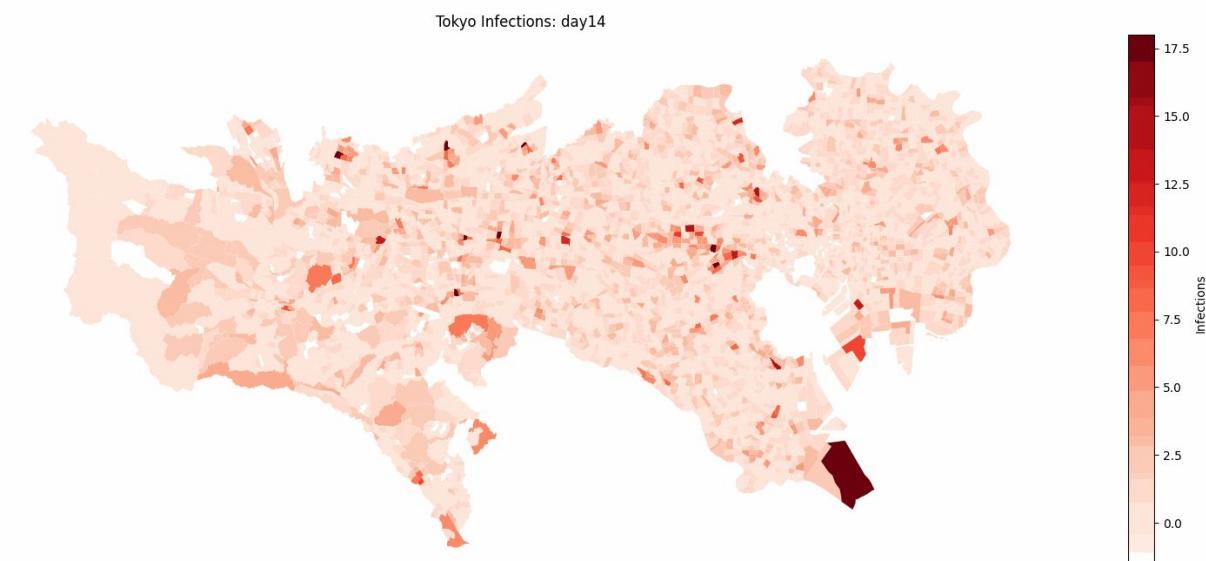
色: ナラティブ

AIエージェントによる見通しの作成、シミュレーション

AIを用いて起こりうる危機の見通しを立て、危機へと備える



AIを用いて、各々のエージェントに個性を付与



感染シミュレーションの例

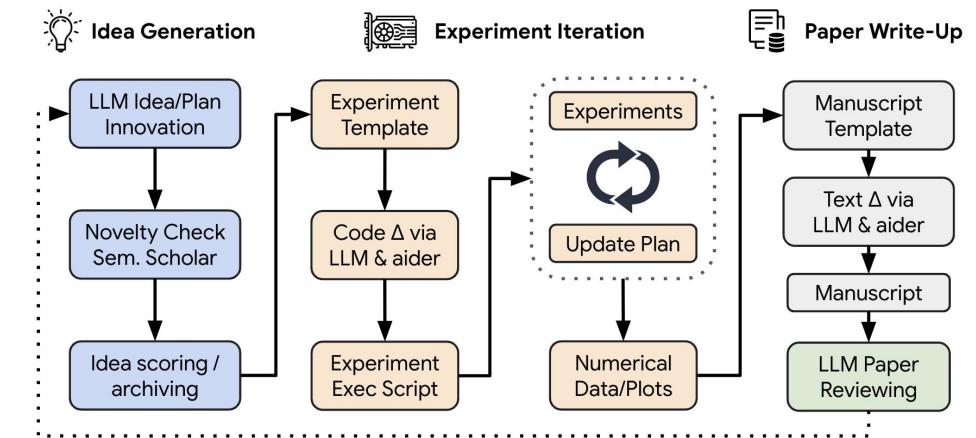
05

おわりに



おわりに

- **Sakana AIとは？**
 - 世界レベルで競争力のある人材による「ユニークなアイデアを探求する会社」
- **防衛分野における取り組み・AI Scientist フレームワークの応用**
 - 無人機の自律的制御への応用
 - インテリジェンス分析への応用
- **他領域への応用**
 - サイバーディフェンスエリアへの取り組み
 - 様々な防衛分野に R&D 成果で防衛産業へ貢献





sakana.ai