

AI研究開発力の強化に向けて

～Science for AI／Science of AI の推進～

(検討)

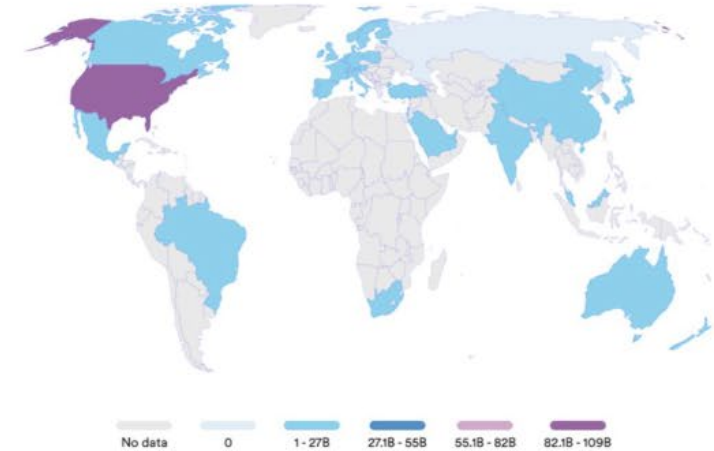
2026年8月7日
研究振興局参事官(情報担当)付

1. 世界のAIに関する情勢

米中の覇権争い（AIに関する複数の指標より）

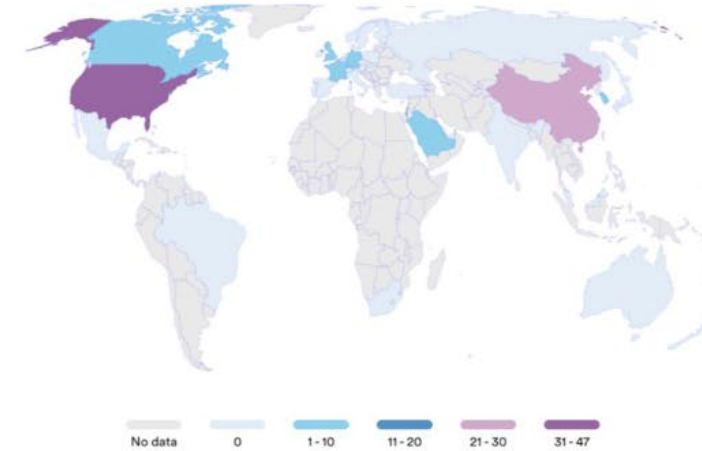
Total AI Private Investment, 2024 **AIへの民間投資**

Source: Quid, 2024 | Chart: AI Index, 2026



Notable AI Models, 2024 **著名なAIモデル数**

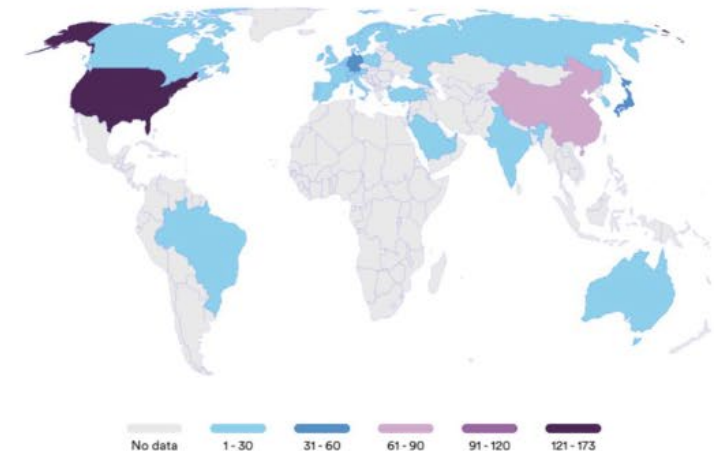
Source: Epoch, 2025 | Chart: AI Index, 2026



**AIへの民間投資、
AIモデル数、計算
資源において米国は
トップであるが、
AIに関する論文生
産では、中国が米国
を追い抜く**

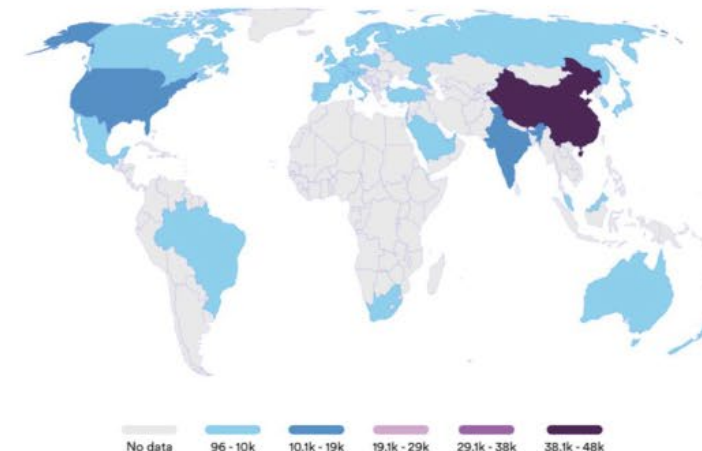
Supercomputers, 2024 **スパコン台数**

Source: Top500, 2024 | Chart: AI Index, 2026



AI Publications (Total), 2024 **論文生産**

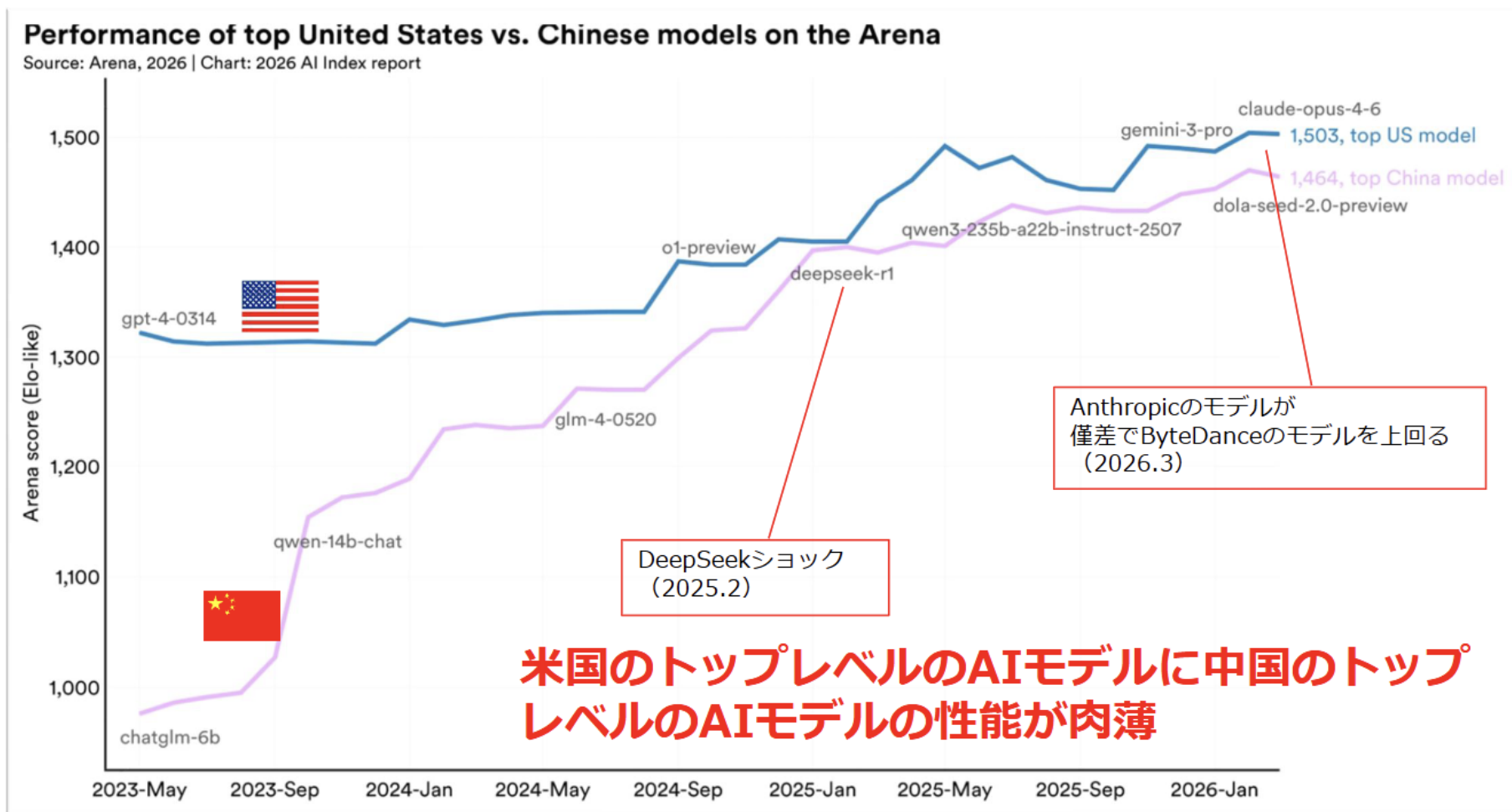
Source: AI Index, 2025 | Chart: AI Index, 2026



Sha Sajadieh, et al., "The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026.
<https://hai.stanford.edu/ai-index/2026-ai-index-report>

1. 世界のAIに関する情勢

米中の覇権争い（トップレベルAIの性能比較より）



Arena score (縦軸) : <https://arena.ai/leaderboard>
ユーザーが2つのAIモデルの出力を比較し優劣を投票。投票結果を集計することにより算出される。

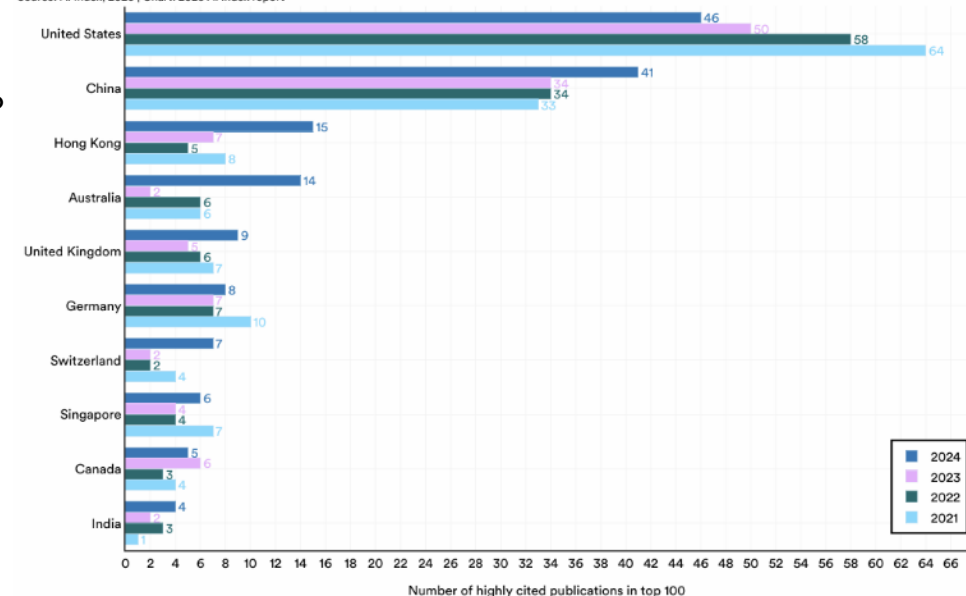
Sha Sajadieh, et al., "The AI Index 2026 Annual Report," AI Index Steering Committee
Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026.
<https://hai.stanford.edu/ai-index/2026-ai-index-report>

(参考)AI Index Report 2026:研究開発分野の主なポイント

1. 2025年には、注目すべきAIモデルの90%超が産業界によって開発された。
一方で、最先端モデルの透明性は低下している。
2. 中国はAI研究で世界をリードし、米国は注目度の高いモデル開発で優位を維持している。
3. 公表されるパラメータ数は数兆規模で横ばいだが、情報開示は減少している。
4. 事前学習において、合成データはまだ実データを代替していないが、データ品質向上や事後学習技術に有望な成果が見られる。
5. 世界のAI計算能力は2022年以降、年平均3.3倍で増加し、H100換算で1,710万基に到達した。
6. AIデータセンターでは米国が圧倒的な優位にあり、その内部で使用するチップの大半は台湾のTSMCが製造している。
7. AIの環境負荷は電力、水資源、排出量のすべてで増大している。
8. オープンソースAI開発は引き続き拡大している。
9. 米国へ移動するAI研究者・開発者の数は、2017年以降89%減少した。
10. AI人材の地理的分布は変化しているが、男女格差は依然として大きい。

Number of highly cited publications in top 100 by select geographic areas, 2021–24

Source: AI Index, 2026 | Chart: 2026 AI Index report



(参考)AI Index Report 2026:科学分野におけるAI活用の主なポイント

1. AI関連の科学論文は年々増加している
2. 分子生物学では、小規模モデルが大規模モデルを上回る成果を示している。
3. 最先端モデルは平均的な人間の化学者を上回るが、公開研究の再現には依然課題がある。
4. 仮想細胞(Virtual Cell)モデルが2025年の新たなフロンティアとして台頭した。
5. 天文学分野では、初の基盤モデル、可視化ベンチマーク、大規模学習データセットが公開された。
6. AIが初めて気象予測プロセス全体をエンドツーエンドで実行した。
7. 2025年、完全にAIが生成した論文が初めて査読付きワークショップに採択された。
8. 科学向けAIモデルの多くは、学術機関や政府機関によって開発されている。

Number of notable AI models by select geographic areas, 2025

Source: Epoch AI, 2026 | Chart: 2026 AI Index report

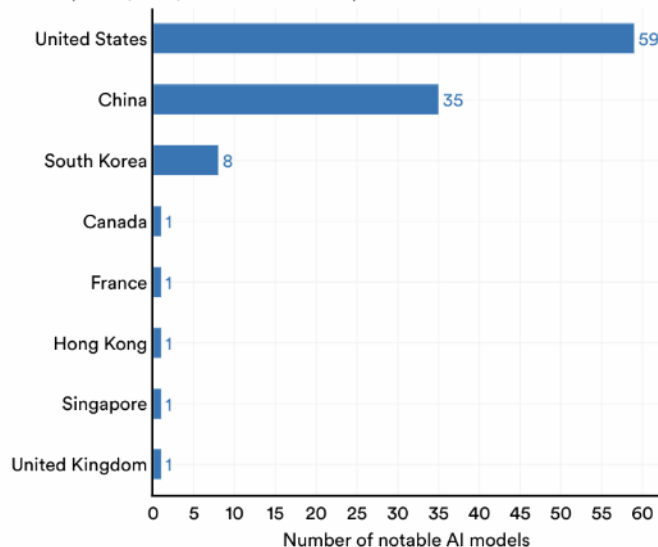


Figure 1.1.1³

Number of top AI authors and inventors by country, 2025

Source: Zeki Data, 2025 | Chart: 2026 AI Index report

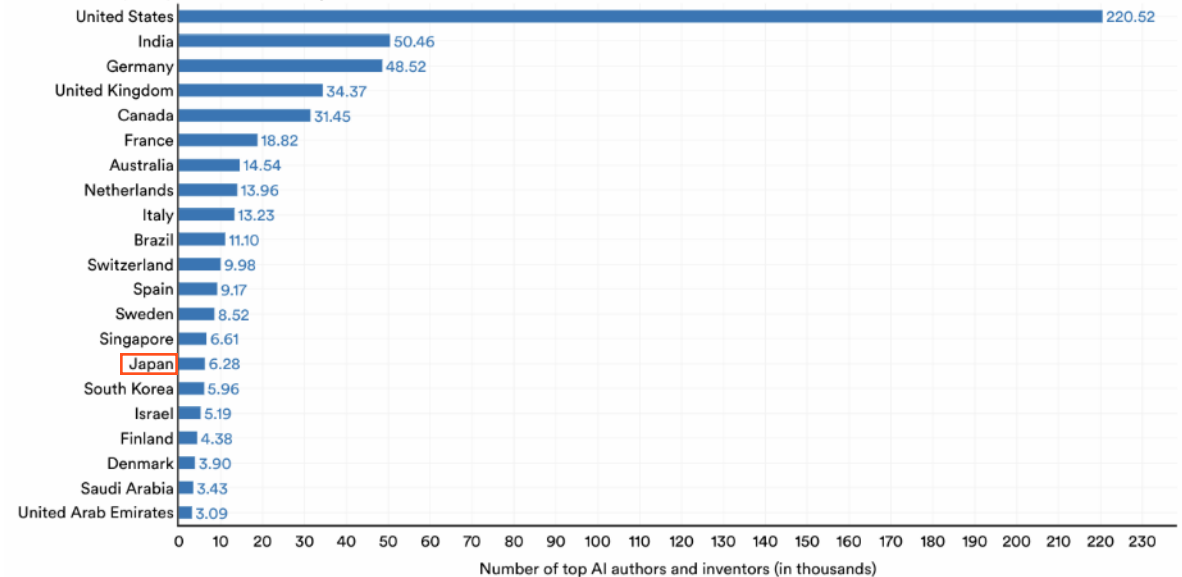
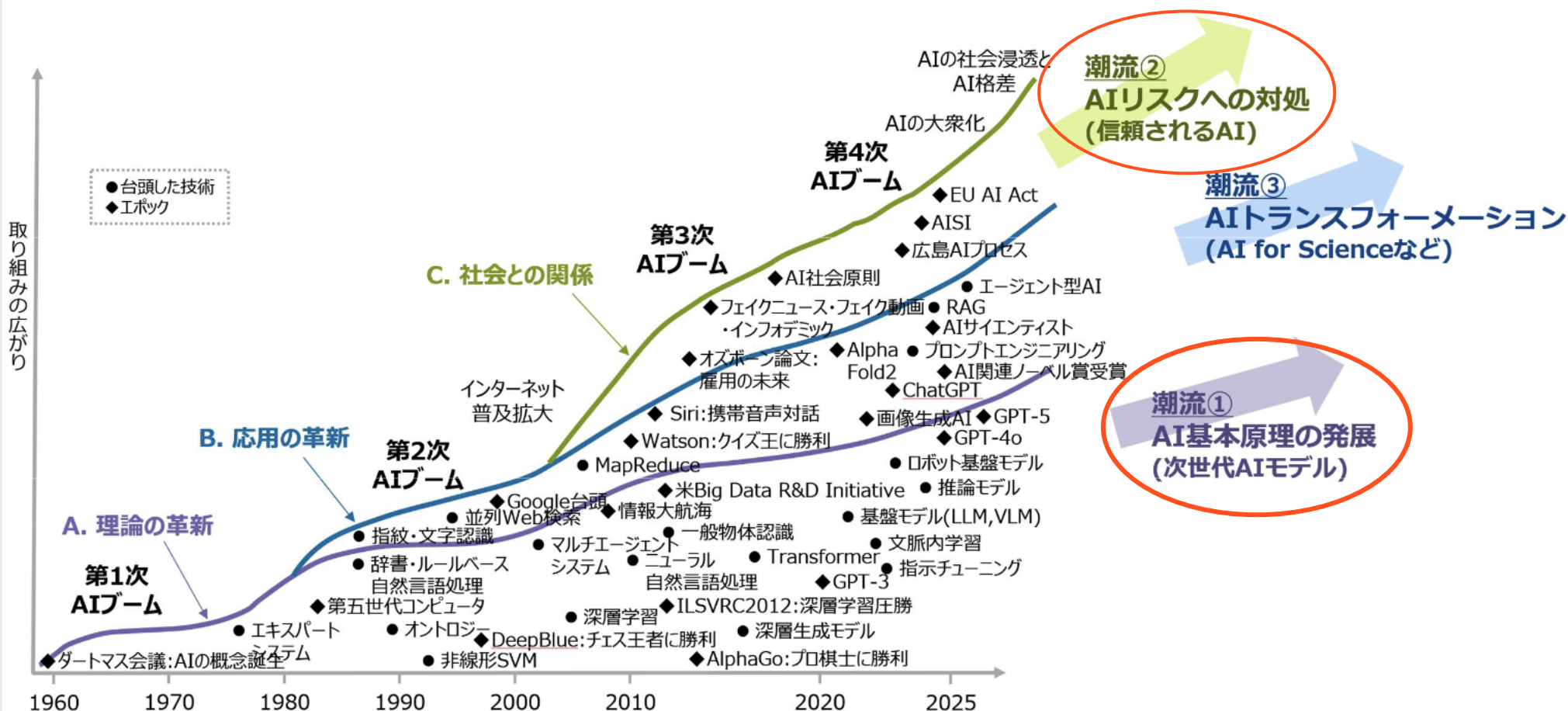


Figure 1.8.1

1. 世界のAIに関する動向

参考：人工知能（AI）の時系列俯瞰図

CRDS 研究開発の俯瞰報告書:システム・情報科学技術分野 (2026年)

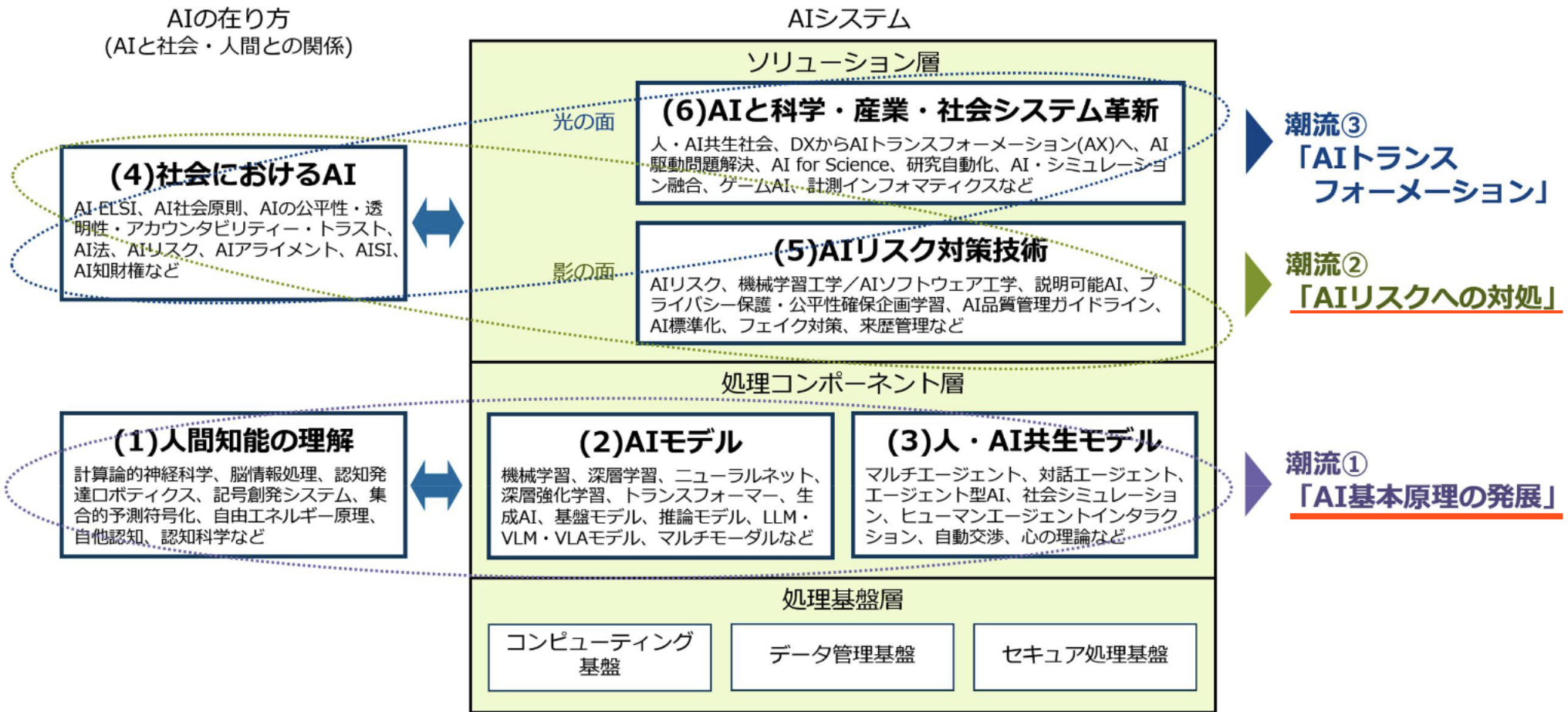
<https://www.jst.go.jp/crds/report/CRDS-FR-S.html>


1. 世界のAIに関する動向

参考：人工知能（AI）の研究開発領域の構造俯瞰図

CRDS 研究開発の俯瞰報告書:システム・情報科学技術分野 (2026年)

<https://www.jst.go.jp/crds/report/CRDS-FR-S.html>



国際動向等

- 生成AIは急速な性能向上を遂げ、科学研究を加速するAI for Scienceが世界的に発展している。一方で、各国政府、国際機関、主要AI企業、第一線研究者の間では、AIを活用する研究(AI for Science)に加え、AIそのものを科学的に理解・高度化する研究が、今後のAI競争力や安全性を左右する重要な研究領域であるとの認識が広がっている。
- 世界の研究開発は、AIを利用するだけでなく、AIそのものを科学の対象として、その原理や能力発現メカニズムを探究する方向へと重点が広がりつつある。

(1) スケーリング競争から「AIの原理」の競争へ(学習原理・アーキテクチャの革新への回帰)

2020年代前半は、大規模モデル、データ及び計算資源を拡大するスケーリング則に基づくAI開発が急速に進展した。一方、近年では、スケーリング則は依然として重要であるものの、それだけでは次世代AIの実現には十分ではないとの認識が広がっている。このため、数学、統計学、情報科学、物理学、脳科学等の知見を活用し、新たなAIの理論・学習原理・アーキテクチャを探究する研究が活発化している。

(2) AIを「ブラックボックス」として利用する段階から、AIを科学的に理解・評価する段階へ

AIの能力向上に伴い、AIシステムの内部メカニズムや能力発現、限界、リスクを科学的に理解し、その挙動を評価・予測・制御可能とする研究が世界共通の課題となっている。International AI Safety Reportでも、AIの能力・リスクを科学的に理解・評価する研究の重要性が指摘されている。

(3) フロンティアAI企業もAIの基礎科学・評価科学への投資を強化

主要AI企業においても、性能競争に加え、AIの原理理解や評価科学への投資が拡大している。例えば、Anthropicは「モデル評価の科学は十分成熟していない」とし、AI評価科学の確立を重要課題として位置付けている。OpenAI、Google DeepMind、Metaなども、推論能力、新たな学習原理、AIの解釈可能性等の基礎研究を強化している。

(参考)海外動向等1/2

国／地域	主な戦略等
米国	• America's AI Action Plan(2025) • The National AI R&D Strategic Plan(2023 Update) • FASST Initiative (DOE, 2024)
中国	• 人工知能+ (AIプラス)行動(2025) • 国民経済・社会発展第15次五カ年計画(綱要)(2026-2030) ※AIを含む国家上位計画
EU	• The AI Continent Action Plan(2025) • AI in Science Strategy(2025)
英国	• AI Opportunities Action Plan(2025) • AI for Science Strategy(2025) • National AI Strategy (2021)
カナダ	• Canada's National Artificial Intelligence Strategy: AI for All(2026) • Pan-Canadian Artificial Intelligence Strategy (PCAIS)(2021～第二期)
ドイツ	• Artificial Intelligence Strategy of the German Federal Government(2020改訂) • German Action Plan Artificial Intelligence (2023) • Hightech Agenda Germany(2025)

(参考)海外動向等2/2

国／地域	主な戦略等
韓国	• National Strategy for Artificial Intelligence (2019) • Korea AI Action Plan(2026－2028)
シンガポール	• Singapore National AI Strategy 2.0(2023)、Update to Singapore’s National AI Strategy(2026) • AI for Science Initiative(2024～) • National AI Research and Development Plan(2026)
フランス	• National Strategy for Artificial Intelligence(2025～第3期) • France 2030(2021～)
スイス	• Digital Switzerland Strategy 2026
オランダ	• Strategic Action Plan for Artificial Intelligence(2019) • Netherlands Digitalisation Strategy(2025更新) • National Technology Strategy (2024)
イスラエル	• National Program for Artificial Intelligence(2026)

- ✓ AI調べ
- ✓ 政府主導型で上位の国家計画、国家AI戦略、AI研究開発計画等が概ねある場合と、いわば研究機関主導型で研究大学・研究機関がAI研究戦略を実質的に担う場合があるため、全てを同等に並べているものではない点に留意が必要。

(参考)米国 国立科学財団(NSF)

■NSF National AI Research Institutes

- NSFは、AIの基礎研究の推進と社会課題の解決、人材育成を目的として、全米の大学・研究機関が連携するネットワーク型AI研究プログラムを2020年に創設。
- AI基礎研究から分野別応用研究、人材育成、産学官連携までを一体的に推進し、AI技術の発展と社会実装を加速する。
- 現在、30を超える研究拠点が全米で活動し、それぞれ異なる重点テーマを担いながら、全米規模のAI研究ネットワークとして運営されている。

■NSF AI Institute for Foundations of Machine Learning

- 機械学習の数理基盤に特化したNational AI Research Institutes第1期の中核研究拠点。AIの持続的な発展に不可欠な機械学習の理論・アルゴリズム・数理基盤を確立し、次世代AIを支える基盤技術を創出することを目的とする。
- 数学的理論と実世界の課題を結び付け、信頼性・効率性・適応性を備えたAIシステムの実現を目指すとともに、理論研究と実践研究の双方を担う次世代AI研究者の育成を推進。



■研究テーマ例

機械学習の基礎理論を中核テーマとし、以下のようなAIの理論・アルゴリズム・学習原理の高度化を推進。

- ・ 機械学習理論
- ・ 汎化理論 (Generalization)
- ・ ロバストネス・敵対的攻撃への耐
- ・ 深層学習アルゴリズム
- ・ 動的・非定常データへの適応学習
- ・ データ効率・計算効率の高いAI
- ・ 最適化理論
- ・ Foundation Modelsを含む大規模機械学習の理論

- 2019年設立。AI技術の急速な発展を踏まえ、人間中心(Human-Centered)の理念の下、AIの基礎研究から社会実装・政策形成までを一体的に推進するため設立。
- 近年は、生成AIの急速な発展を受け、Foundation ModelsやAI Agentsなど次世代AIの基盤研究に加え、生命科学・医学等へのAI活用(AI for Science)やAIガバナンス・安全性に関する研究を重点的に推進。
- AI研究だけでなく、教育、人材育成、産学連携、政策提言を通じて、人間中心のAIEcosystem全体を牽引する世界的研究拠点を目指す。

■取組事例

(1) Science of AI

- Foundation Models
- Large Language Models
- AI Agents
- Spatial Intelligence
- Reasoning
- Efficient AI
- World Models

(2) AI for Science

- Biology
- Medicine
- Healthcare
- Earth Science
- Climate
- Scientific
- Discovery

(3) Human-centered AI

- AI Ethics
- AI Policy
- Responsible AI
- Human-AI Interaction
- Democracy

■概要(特徴)

- 人間中心(Human-Centered)の理念の下、AI基盤研究、AI for Science、社会実装、政策形成を統合的に推進する大学横断型AI研究拠点。
- 工学・医学・自然科学・社会科学・人文学を横断する学際的研究と、教育・産学連携・政策提言を一体的に推進することが特徴。
- Google、Microsoft、NVIDIA、Salesforce、OpenAI、他スタートアップ多数参画。



(参考) 英国 2つのAI研究ラボを開設

- 英国研究・イノベーション機構(UKRI)は、英国のAI研究・イノベーションを強化するため「UKRI AI Research and Innovation Strategic Framework(※)」を策定(2026年6月、4年間総額16億ポンド)。
- 現在のAIの限界を超える次世代AIの実現に向けて、AIの基盤となる基礎科学を推進するネットワーク型の研究拠点(AI Research Labs)の整備を開始(6年間で最大6千万ポンド)。

■SOFAIR Lab

(Science of Fundamental AI Research Lab
:基礎AI科学研究ラボ)



【参画大学】

ユニヴァーシティ・カレッジ・ロンドン(UCL)を筆頭に、ケンブリッジ大学、オックスフォード大学、エディンバラ大学が協働。

【取組内容】

現在のAIシステムが抱える「ハルシネーション」や「特定企業への依存」といった根本的課題を解決するため、コンピュータ科学、数学、統計学、神経科学を融合させ、現行Transformerを超える新しいAIアーキテクチャ、低計算資源で動くAI、AIをよりアクセスしやすくする基盤技術、オープンなAI基盤技術の設計・理論構築を探究する研究を推進。

■BOLD Lab



(British Open-ended Learning and Discovery Lab:英国オープンエンド型学習・発見ラボ)

【参画大学】

オックスフォード大学を筆頭に、ユニヴァーシティ・カレッジ・ロンドン、インペリアル・カレッジ・ロンドンが協働。

【取組内容】

巨大テック企業のように膨大なデータと中央集権的な計算資源の投入によりスケールアップさせる従来手法とは異なる、AIが「より効率的に、自律的に、環境に適応して学習する」ためのまったく新しい学習方法や新しいAI学習理論、次世代AIアルゴリズムなどの開発を推進。AIが環境との相互作用を通じて継続的に学習・発見し、人間の共同研究者(co-scientist)として科学的発見を支援する次世代AIの実現を目指す。

(※) UKRI AI戦略:主な項目は以下の通り

①AIの技術開発と未来の基盤、②AIによる研究の変革、③AIスキルと人材の育成、④経済成長と社会的利益のためのイノベーションと採用の加速、⑤責任ある信頼できるAIの提唱、⑥世界クラスのAIデータおよびインフラストラクチャの構築。

- 欧州のAI研究力の強化と世界トップレベルのAI研究・人材の維持を目的として、欧州の機械学習研究者が中心となり2018年に設立。
- 北米・中国へのAI研究人材流出や研究基盤の分散といった課題を踏まえ、欧州全体を結ぶ分散型研究ネットワークを構築し、世界最高水準のAI基礎研究・人材育成を推進。
- AIの理論・アルゴリズムなどの基礎研究から、ロボティクス・医療・気候科学等への応用までを欧州全体で推進することを目指す。
- ELLIS PhD Programでは、複数国・複数大学の研究者による共同指導を行い、欧州全体でAI研究者を育成。「研究者育成」を組織の中心ミッションに据えている。

■取組事例

AIの基礎科学を中核テーマとして、

- 機械学習理論
- Deep Learning
- 強化学習
- 因果推論
- 確率的機械学習
- コンピュータビジョン
- 自然言語処理
- AI Safety・Trustworthy AI
- Resource-efficient AI
- Human-centered AI

など、AIの理論・アルゴリズム・安全性・効率性に関する基礎研究を推進

■概要

- AI理論・機械学習の基礎研究と人材育成に特化した欧州の分散型研究ネットワーク。
- 約40のELLIS Units(研究拠点)と約16のResearch Programsにより構成され、欧州17か国以上の大学・研究機関が参画。



(参考) フランス AISSAI(AI for Science, Science for AI Centre)

- フランス国立科学研究センター(CNRS)は、AIが科学研究やAIそのものの発展を大きく変革することを踏まえ、AI for ScienceとScience for AIを一体的に推進する中核拠点としてAISSAIを設立(2021年設立、2024年にCNRSの正式研究支援組織(UAR)として運用開始)。
- AI研究者と自然科学・人文社会科学研究者の協働を促進し、AIを活用した科学的発見の加速と、数学・統計学・情報科学等の知見を活かした次世代AIの創出を目指す。
- CNRS全体のAI戦略を牽引するハブとして、研究、人材育成、国際連携を推進。
- AI研究者だけでなく、各分野の研究者がAIを活用・共同開発できる人材育成を推進。

■ Science for AI

数学・統計学・物理学などの知見を活用し、

- 機械学習理論
- Explainable AI
- Trustworthy AI
- Efficient / Frugal AI(計算効率の高いAI)
- 専門知識を取り込むAI
- Scientific Machine Learning
- 因果推論
- Scientific Foundation Models

など、AIの理論・アルゴリズム・学習手法の高度化を推進。

■ 概要

- CNRS(フランス国立科学研究センター)の全分野を横断するAI研究支援・連携拠点。
- AI for ScienceとScience for AIを一体的に推進し、AI研究者と各分野の研究者の協働を促進する全国横断型のハブとして、研究、人材育成、国際連携を推進。
- Science for AIを組織名称に明示的に掲げる世界でも数少ない研究拠点。



(参考) スイス ETH AI Center

- スイス連邦工科大学チューリッヒ校(ETH Zurich)は、AIが科学・産業・社会全体を変革する基盤技術となることを踏まえ、大学全体のAI研究・教育・イノベーションを統合する中核拠点としてETH AI Centerを2020年に設立(2021年から本格始動)。
- AIの基盤研究、各分野への応用研究、社会・倫理・政策を一体的に推進し、「信頼できる(Trustworthy)、包摂的(Inclusive)、誰もが利用できる(Accessible)AI」の実現を目指す。
- 学際的研究、産学連携、人材育成、スタートアップ創出を通じて、AI研究から社会実装までを一体的に推進。AI研究者だけでなく、各専門分野でAIを活用・開発できる人材を育成。

■取組事例

AIの理論・方法論の高度化をはじめ、

- 機械学習
- Foundation Models
- 深層学習
- コンピュータビジョン
- 自然言語処理
- AI Safety
- Trustworthy AI
- Explainable AI

など、AIの理論・アルゴリズム・信頼性に関する基盤研究を推進。

■概要

- 大学全体のAI研究・教育・社会実装を統合する総合AIセンター。
- AI Foundations(AI基盤研究)とAI Applications(AI応用研究)が相互に発展する(co-evolve)研究モデルを採用し、学際研究・産学連携・スタートアップ創出までを一体的に推進することが特徴。



ETH AI CENTER

(参考) 中国 新世代人工知能国家科学技術重大プロジェクト

- 2026年度に、新一代人工智能重大科技专项(大型研究開発プロジェクト)として第1回公募を開始。
- 次世代AIの中核技術の突破を図り、AI分野における国家戦略能力、科学技術競争力及び自主イノベーション能力を強化するとともに、国家のAI戦略と科学技術自立を支える研究開発能力の向上を目指すもの。
- AI基礎理論、基盤モデル、AI Agent、具身知能(Embodied AI)、AI for Science等の重点研究課題を推進する国家最上位級のAI研究開発プロジェクトであり、2017年「次世代人工知能発展計画」等を踏まえた国家AI研究開発プログラムとして位置付けられる。

■概要

- ・ 中国科学技術部(MOST)の下で実施される国家最上位級のAI研究開発プロジェクト。
- ・ 大学、研究機関、企業等によるコンソーシアムを対象に、AI基礎理論、基盤モデル、AIエージェント、具身知能、AI for Science等の重点課題を3～5年程度の大型共同研究として推進する。
- ・ 1プロジェクト当たり複数課題(原則5件以内)、参加機関10機関以内で実施され、課題ごとに研究費が設定される。

■研究テーマ例

- | | |
|----------------------------|------------------------|
| ・ AI基礎理論 | ・ AI for Science |
| ・ 基盤モデル(Foundation Models) | ・ AI Computing |
| ・ AI Agent | ・ AI Software・Hardware |
| ・ Embodied AI(具身知能) | ・ 科学・産業分野へのAI応用 など |

目指す世界像

急速なAI技術の進展に伴い、以下のような課題への対応が求められる。

- モデルの透明性・安全性・信頼性の向上、高品質データの確保
- 計算資源・推論コスト・消費電力の増大への対応
- 経済安全保障も踏まえたAI分野の戦略的自律性の確保



- AI for Scienceに加え、
- ・ 科学的知見を活用して新たなAI技術を創出する「Science for AI」、
 - ・ AIの原理や能力・限界、安全性・信頼性を科学的に解明する「Science of AI」
- において、日本が国際的に不可欠な知的基盤となることが重要。



AI for Scienceの推進と相互に補完・強化しながら、「Science for AI」及び「Science of AI」を推進することで、世界のAIと科学の進化に貢献する国を目指す。

日本の現状(強みと課題)

強み

- 日本は、AIそのものの基盤となる数学、統計科学、情報科学、物理学、脳科学、ロボティクス等の基礎研究分野において、一定の研究蓄積と人材を有してきた。
- ロボティクス・身体性AI、脳科学・認知科学、科学計算・シミュレーションなど一部の領域は、日本が国際的に競争力を維持する数少ない領域であり、Science for AI/Science of AIを推進する上での強みとなり得る。
- 日本企業においても、製造業、ロボティクス、材料、医療等の分野で、高品質なデータ、専門知識、現場技術を蓄積しており、専門知識とAIを融合する領域において優位性を発揮できる可能性がある。

課題

- 現在のAI競争においては、単なるAI活用能力ではなく、AIの原理を理解し、次世代AIを自ら設計・高度化する能力が重要になっている。しかし、日本では以下の課題が存在。

(1)AI基礎研究を国家戦略として推進する枠組みが弱い

- 米国や英国、中国などでは、AIの基礎理論、新たな学習原理、次世代AIアーキテクチャを対象とした国家的研究投資が進んでいる。一方、日本では現在、AI活用・社会実装に重点が置かれており、AIそのものの原理解明や次世代AI創出を目的とした長期的・基礎的研究プログラムや研究拠点形成は、国際的に見て限定的である。

(2)研究力は存在するが、分散しており、国際的拠点形成につながっていない

- 日本には、個別に高い研究力を持つ拠点が存在する。しかし、AI研究者、数学・物理・生命科学等の研究者、産業界、若手研究者育成を結集し、次世代AIの基礎研究を継続的に推進する国際的な研究拠点・ネットワーク型モデルは十分に形成されていない。

(3)AIを「使う人材」だけでなく「創る人材」の育成が必要

- 生成AIの普及により、AIを活用できる人材育成は進展している。一方、今後必要となるのは、AIの学習原理を理解する研究者、新しいAIアルゴリズムを創出する研究者、AI安全性・評価科学を担う研究者、科学分野とAIを融合できる研究者である。特に、数学、情報科学、自然科学を横断して次世代AIを研究できる人材は国際的にも不足しており、長期的な育成戦略が必要である。

今後の検討課題等

Science for AI／Science of AI の観点から…

日本の強みや特色を踏まえ、

1. 日本(国)として力を入れるべき研究領域・テーマは何か？

- ① 自律性等から取り組むべき課題は何か？
- ② 日本の強みや特色を生かして、世界を先導し得る課題は何か？
- ③ 産業界や社会で活用される技術や研究は何か？

2. 当面の課題と中長期的な方向性は何か？

(人材育成、研究基盤、産学連携、国際連携等を含め、どのようなエコシステムを構築すべきか)

■必要な機能等の検討

- ① 推進体制(ガバナンス体制、役割分担や明確化、全国的なネットワーク型研究体制など)
 - ② 支援のあり方(中長期・安定的な研究支援など)
 - ③ 人材育成の考え方(博士課程学生・若手研究者の育成など)
 - ④ 産学連携の考え方(企業との共同研究・社会実装など)
 - ⑤ 国際連携の考え方(国際共同研究・研究者交流など)
 - ⑥ AI for Science との一体的推進
- など

(参考資料)

(参考) 大まかな概念イメージ (案)

AI for Science (科学のためのAI)	<u>AIを活用して、科学研究を加速・変革する取組</u> - 実験・観測・シミュレーション等から得られる膨大なデータをAIで解析し、仮説生成、実験設計、予測、知識発見等を高度化・高速化することで、科学的発見や社会課題解決を加速。 - データ、計算資源、研究設備等を統合したAI駆動型研究の基盤(研究インフラ)整備が重要。
Science for AI (AIのための科学)	<u>科学の知見を活用して、AIを高度化する取組</u> - 数学・統計学・情報科学・物理学・脳科学・認知科学等の知見を活用し、学習効率、推論能力、汎化性能、計算効率等を向上させ、次世代AI技術を創出。 - AIの基盤技術は、国の競争力及び経済安全保障の中核を成す戦略技術であり、中長期的な基礎研究・人材育成への投資が不可欠。
Science of AI (AIの科学)	<u>AIそのものを科学的に理解・解明する取組</u> ※国際的に確立した定義ではない点に留意が必要 - AIシステムの能力、限界及び創発現象等を理解し、その挙動を評価・予測・制御可能なものとするを通じて、安全性、信頼性、透明性及び説明可能性を向上。 - 近年、AIの高度化・複雑化に伴い、AIの理解・評価・制御可能性を高める研究の重要性が高まっている。 例: 米国ではAIアクションプラン(2025年7月)において“Advance the Science of AI”や“Interpretability, Control, and Robustness”等が明記されている。また、英国王立協会(2024年5月)も、AIのブラックボックス性が科学研究の検証可能性や再現性、信頼性の課題となり得ることを指摘している。

「Science of AI」はAIの内部機構や能力・限界を科学的に解明し、「Science for AI」による次世代AI技術の創出や、「AI for Science」で活用されるAIの安全性・信頼性向上の基盤となる。高性能AIによって加速される科学的発見から得られる新たな知見や課題は、三区分別々の発展を促す。

➤ この三者の好循環を戦略的に推進することで、AIと科学技術の双方の飛躍的发展が期待される。

(参考)主な関係ワード例

Science for AI (AIのための科学)	Science of AI (AIの科学)
(1)AIの基盤理論・基盤技術 • Foundation Models(基盤モデル) • Machine Learning Theory(機械学習理論) • Efficient AI(高効率AI) • Scientific Machine Learning(科学機械学習) • AI Hardware(AIハードウェア・アーキテクチャ) • Frontier AI Models(フロンティアAI) • Multimodal AI(マルチモーダルAI) • World Models(ワールドモデル)※環境・世界の構造や時間変化を内部モデルとして学習・予測するAI • Scaling Laws(スケーリング則) • Inference-time Scaling(推論時スケーリング) など (2)AIの高度化・次世代AI • Agentic AI(エージェント型AI) • Embodied AI(身体性AI) • Physical AI(フィジカルAI) • AI Scientist(AIサイエンティスト) • Knowledge-integrated AI(知識統合型AI) • Scientific Foundation Models(科学基盤モデル) • Autonomous Scientific Discovery(自律的科学発見) など	両者は密接に関係しており、境界が明確ではない点に留意が必要。 (3)AIの理解・解明 • Theory of Deep Learning(深層学習理論) • AI Reasoning(AI推論能力) • Interpretability(解釈可能性) • Explainable AI(説明可能なAI(XAI)) • Brain-inspired AI(脳型AI) • AGI Theory(AGI理論(汎用人工知能理論)) など (4)AIの信頼性確保 • Trustworthy AI(信頼できるAI) • AI Safety(AI安全性) • Alignment(アラインメント) • Evaluation(AI評価) • Robustness(ロバスト性) • Verification(検証) • Validation(妥当性確認) • AI Security(AIセキュリティ) など (5)AIと人間・社会 • Human-centered AI(人間中心AI) • AI Governance(AIガバナンス) • Responsible AI(責任あるAI) • Human-AI Interaction(人間とAIの相互作用) • AI Ethics(AI倫理) • AI Risk Science(AIリスク科学) など

(参考) 米国 America's AI Action Plan(2025)ポイント

■ Advance the Science of AI(AIの科学の進歩)

AIの能力を前進させ、新たな変革をもたらすパラダイムに向けて、AIシステム的能力、効率性、信頼性を向上させる理論的(theoretical)、計算論的(computational)、実験的(experimental)なブレークスルーにつなげる研究への投資を優先する。これには、現在のトランスフォーマーアーキテクチャの限界を超える可能性のある、「エネルギー効率の高い代替方式」や「神経形態学的(ニューロモーフィック)コンピューティング」などの次世代アーキテクチャの研究開発(R&D)の支援が含まれる。

次期国家AI研究開発戦略(The next National AI R&D Strategic Plan)においては、AI研究における根本的な進展を優先事項とすべきである。

■ Invest in AI Interpretability, Control, and Robustness Breakthroughs

(AIの解釈可能性、制御、および堅牢性の画期的な進歩への投資)

フロンティアAIシステムがどのように意思決定を行うかを人間が理解・測定・監視できるようにするためのツールの開発(Interpretability)、AIシステムが人間の意図や明確に定義された境界条件から逸脱しないように確保するための制御メカニズムの開発(Control)、データ汚染や敵対的攻撃といった脆弱性に対する堅牢性の向上(Robustness)を進める。

■ Build an AI Evaluations Ecosystem(AI評価エコシステムの構築)

AI標準・イノベーションセンター(Center for AI Standards and Innovation:CAISI)等の体制を通じてフロンティアAIモデルの評価を行う。研究者が安全な環境でAIモデルのリスクや性能を実証・試験できる「安全なテストベッド」を構築・提供するとともに、AIモデルの性能、信頼性、能力、脆弱性、バイアス及び国家安全保障上のリスクについて、科学的・客観的な評価・測定を可能とする共通基準を策定する。²⁸

政策文書等における主な関係記載(抜粋)

■経済財政運営と改革の基本方針2026(令和8年7月、閣議決定)

- ・改定「人工知能基本計画」17等を踏まえ、リスクの適正な管理の下、AI開発・利活用を強化し、開かれたAI主権を 確立しながらAI・データ基盤を整備する。
- ・AIを科学研究 に適用し、新たな発見や理解を加速する「AI for Science」を推進する。

■日本成長戦略(令和8年7月、閣議決定)

- ・重点領域の領域別戦略を策定し、人材を含めAI基盤や研究開発・実証への集中投資と、政府調達や制度改革による初期需要創出を一体的に推進する。
- ・主な施策:パーティカルAIの導入・開発・利用を支えるAI基盤の整備、研究開発、実証、データ創出を進める。

■パーティカルAI領域別戦略中間とりまとめ(令和8年7月、人工知能戦略本部)

③課題解決の手段・アプローチ

- ・科学研究に安全に活用可能な、信頼できる(される)AIの開発に向けたAIの基礎研究を含むAIそのものの研究を強化する。

④2030年に向けて取り組む主な施策

世界を先導する研究成果及びイノベーションの創出

- ・日本の強みを活かした最先端研究開発へのAI利活用を加速するとともに、AI自体の基礎研究の推進により、世界を先導する研究成果を創出する。

■統合イノベーション戦略2026(令和8年7月、閣議決定)

- ・AI研究(Science for AI)においては、数理科学を基盤としたAI技術に関連する最先端の理論研究やアルゴリズム開発等の基礎研究を推進するとともに、リスクや懸念の克服に向け、テキスト、画像、音声等をコンテキストや非言語情報も踏まえて処理できるマルチモーダルモデルを含めたAI基盤モデル等の透明性・信頼性の確保及び高度化や、質の高いデータの整備に向けた研究開発を推進する。

■第7期科学技術・イノベーション基本計画(令和8年3月、閣議決定)

(1)AI利活用研究(AI for Science)とAI研究(Science for AI)の推進

■人工知能基本計画(令和8年7月改定、閣議決定)

第2節 AI開発力の戦略的強化

(1) AI研究開発・利用基盤の拡充・高度化

■自民党 AIホワイトペーパー2026(令和8年5月)

- ・科学研究において重要なのは一般的な性能の高さだけではなく、安全性、信頼性、再現性及び継続運用性を備えた形でAIを実装することであり、「規制の強弱」から「信頼の設計」への展開もまた不可欠である。

第Ⅱ期人工知能基本計画（概要）

基本構想

- ◎ 急伸する自律行動型 A I に適応し、計算資源から制度まで日本の A I 実装能力を抜本強化
- ◎ A I を軸に組織や社会を変革する「A I トランスフォーメーション」(A X) を推進
- ◎ **信頼できる A I** で社会全体を駆動し、課題解決と競争力強化を実現

4つの原則

- イノベーション促進とリスク対応の両立
- 挑戦と学習
- アジャイル（柔軟かつ迅速）な対応
- 内外一体での政策推進

4つの基本的な方針に基づく施策

政府一丸でAX推進に必要な体制を構築、あらゆる制度や運用を先導見直し

1. A I 利活用の加速的推進「A I を使う」：自律行動型 A I 等の積極適用、政府における率先導入

- 政府・自治体での A I の徹底した利活用 ガバメント A I 「源内」の進化と普及、地方の伸び代を成長に転換（地域 A X）
- A I 利活用促進による社会課題解決、新事業創出 バーティカル A I ・フィジカル A I 中心に社会実装加速
- データ戦略、更なる A I 利活用に向けた仕組みづくり 政府・準公共・産業分野でデータ連携基盤構築

2. A I 開発力の戦略的強化「A I を創る」：A I 実装能力を高め、独自の価値を創出

- A I 研究開発・利用基盤の拡充・高度化 国家基盤として計算資源やデータ、安定した電力供給を確保
- 日本国内の **A I 開発力の強化** データセット創出、内外トップ人材確保、モデル開発高度化、評価基盤・テストベッド整備
- 信頼できる A I 基盤モデル等の開発 質の高い日本語データによる A I 開発・評価、マルチモーダル基盤モデル開発推進
- 日本の勝ち筋の追求 バーティカル A I （領域別戦略）、フィジカル A I （A I ロボティクス、自動運転）、AI for Science

利活用と
技術革新
の好循環

3. A I ガバナンスの主導「A I の信頼性を高める」：「信頼」を再現するため、制度・運用の機動的見直し

- 責任あるアジャイル・ガバナンスの実現 AI法を含めた制度等の能動的かつ不断の見直し、Project YATA-Shield、青少年保護
- A S E A N 等グローバルサウス諸国を含めた国際協調 AISI機能強化、信頼できる A I 共創に向けた A I サミット早期日本開催

4. A I 社会に向けた継続的変革「A I と協働する」：A X で日本を駆動、自律行動型 A I を前提の社会構築

- A I を前提とした産業構造・雇用環境の構築 中小企業、地方含む A X 推進（取組可視化等）、雇用への影響を調査研究
- A I 社会における制度・枠組みの検討・実証 規制・制度改革、人と A I の責任関係、知財保護・利活用につながる透明性確保
- A I 時代の **人材の育成・確保** A I 実装人材育成・確保、全ての国民の A I に関するリテラシー向上、リ・スキリング支援
- A I 時代における人間力の向上 判断力、批判的思考など人間力向上、人と A I の役割分担の探究

➡ 本基本計画は当面毎年変更、適切なベンチマーク（KPI等）を設定しモニタリング

AI for Scienceの推進に向けた基本的な戦略方針(抜粋)

1-3. 喫緊の課題

そのため、AIを活用した科学研究およびAIそのものの研究開発への取組を抜本的に増強するとともに、産学官連携を通じた科学とビジネスの好循環を創出し、AIイノベーションを持続的に推進していくことが不可欠である。

4. 具体的な目標と手段

加えて、AIの基礎研究を含むAIそのものの研究の強化により、信頼できる（される）AI技術の確保・発展を図る。

7. 基本的な各取組の考え方

加えて、AIは次世代の社会インフラであり、他国に依存することなく、自国で研究開発する能力を保持することは安全保障上も極めて重要であり、信頼できる（される）AIに関する取組や、AIそのものの研究開発を持続的に強化すべきである。

（1）研究：AI研究・AI利活用研究における先駆的・先導的取組の推進

AI研究（Science for AI）においては、数理科学を基盤としたAI技術に関連する最先端の理論研究やアルゴリズム開発等の基礎研究を推進するとともに、リスクや懸念の克服に向け、テキスト、画像、音声等をコンテキストや非言語情報も踏まえて処理できるマルチモーダルモデルを含めたAI基盤モデル等の、透明性・信頼性の確保及び高度化や、質の高いデータの整備に向けた研究開発を推進する。

（2）人材：AI関連人材の育成・確保

産学の研究者が知見や経験を共有する拠点を形成するとともに、AIやデータ等に関するスキルを持った高度専門人材等を幅広い分野で育成・確保し、多様な研究者のAI開発・活用ニーズに対応できる支援体制の構築を通じて、国内におけるAI研究開発力の底上げとAIを高度に利活用可能な研究開発人材の育成を推進する。

おわりに

すなわち、AI for Scienceを前提とした研究システムを構築するとともに、Science for AIを両輪として強化し、国際的な研究エコシステムにおいて不可欠な役割を担う存在となることを目指す。

（別添1）アクション項目

- ⑤ 科学研究に安全に活用可能な、信頼できる（される）AIの開発に向けて、AIの基礎研究を含むAIそのものの研究を強化する。

(参考)システム・情報科学技術分野～分野別俯瞰編～(2026 年)【JST】抜粋

- 研究開発の現状の全体像を把握・分析・可視化することに加え、CRDSが考える今後のあるべき方向性・展望を顕在化させるため、四つの考え方に基づいて国として推進すべき10の研究開発課題を抽出。

表 国として重点的に取り組むべき10の研究開発課題と四つのシナリオ

カテゴリー	研究開発課題	シナリオ*			
		技術	産業	社会	基盤
デジタル安全保障に対する総合知による取り組み	AI 社会のリスク対策・トラスト基盤	●		●	●
	コグニティブセキュリティ				●
	AI トランスフォーメーション基盤		●		●
スマート化・自律化の根本である知能の原理探求	次世代 AI モデル			●	●
	フィジカル AI システム	●	●	●	
	バイオハイブリッドロボット	●	●		
	因果推論	●			●
	最適化	●			●
サステナブル社会のための ICT 基盤	通信と計算の融合	●			●
	AI コンピューティング		●		●

* 技術：強い技術を核とした骨太化、産業：強い産業の発展・革新の推進、社会：社会受容性の活用による社会課題の先行解決、
基盤：経済安全保障上の重要基盤技術の確保。