

2 | 俯瞰区分と研究開発領域

2.1 人工知能・ビッグデータ

人工知能（AI：Artificial Intelligence）技術は、人間の知的活動（認識、判断、計画、学習など）をコンピューターで実現するための技術群である。AI研究として、人間の知能のさまざまな側面を広くカバーし、さまざまな状況で人間の知能のように動作する汎用性の高いシステム（汎用AI）を目指す取り組みがある一方、特定の機能や特定の状況下でのみ人間に近い（時には精度で人間を上回る）振る舞いをするシステム（特化型AI）の開発が活発に進められている。現在の第3次AIブームにおいてさまざまな応用に広がったAI技術は、基本的に特化型AIに相当する技術群であるが、極めて大規模な学習によって作られた基盤モデル（Foundation Model）が登場し、AIの汎用性・マルチモーダル性が急速に高まりつつある。

一方、ビッグデータ（Big Data）は、元来は膨大な量のデータそのものを指す言葉だが、その収集・蓄積・解析技術は、大規模性だけでなくヘテロ性・不確実性・時系列性・リアルタイム性などにも対応できる技術として発展している。また、センサー、IoT（Internet of Things）デバイスの高度化と普及によって、さまざまな場面で実世界ビッグデータが得られるようになり、その収集・解析技術は、実世界で起きる現象・活動の状況を精緻かつリアルタイムに把握・予測するための技術としても期待されている。今日、さまざまな社会課題が人間の手に負えないほどに大規模複雑化しており、実世界ビッグデータの収集・解析による状況の把握・予測は、そのような課題の解決に共通的に貢献し得る有効な手段になる。

これらAI技術とビッグデータ（データそのもの、および、処理技術）は深く関係し合いながら発展している。ビッグデータが集められることでAI技術（特に機械学習技術）は高度化し、精度を高め、そのAI技術を用いて実世界のビッグデータを解析することで、実世界の現象・活動のより深く正確な状況把握・予測が可能になってきた。

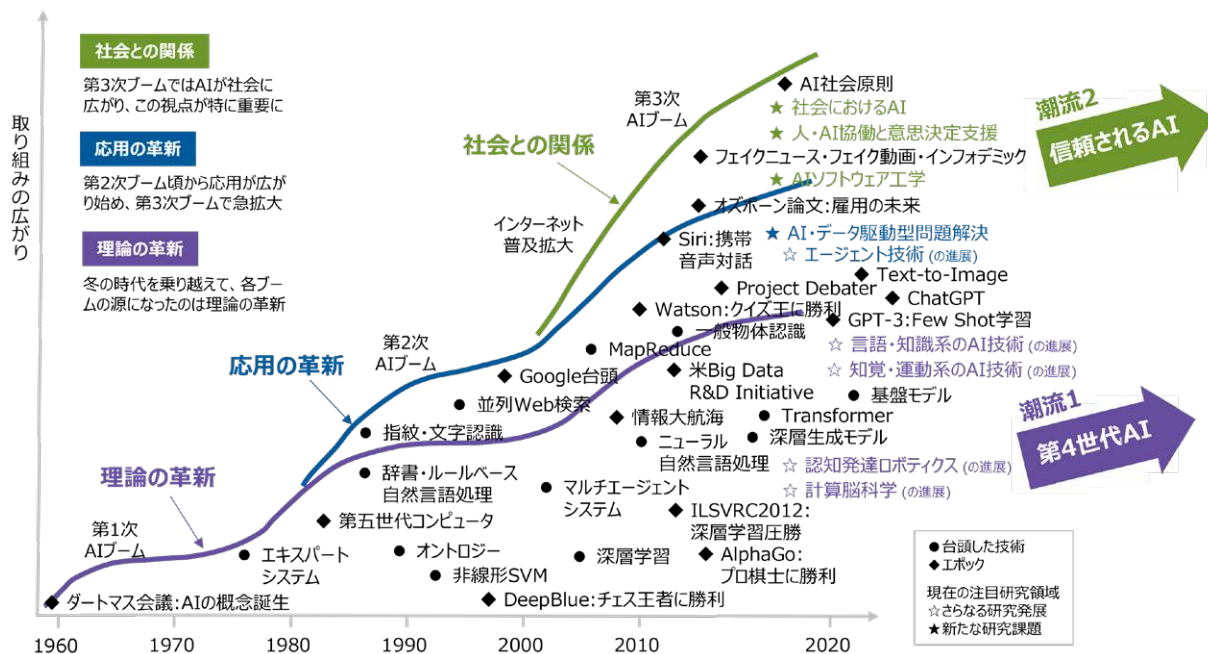


図 2-1-1 人工知能・ビッグデータ技術の俯瞰図（時系列）

【AI・ビッグデータの俯瞰図（時系列）】

AI・ビッグデータの技術発展に関する俯瞰図（時系列）を図2-1-1に示す。この図では、横軸が年代、縦軸が取り組みの広がりをおおまかに表している。図中には、その時期に台頭した技術およびエポックをプロットした。AIは現在、3回目のブームを迎えているが、これまでブームと連動して取り組みが広がってきた様子を示している。

第1次AIブーム（1950年代後半から1960年代）では、AIに関わる基礎的な概念が提案され、AIが新しい学問分野として立ち上がったが、実用性という面ではまだトイシステムであった。第2次AIブーム（1980年代）では、人手で辞書・ルールを構築・活用するアプローチが主流となり、エキスパートシステム、指紋・文字認識、辞書・ルールベース自然言語処理など（カナ漢字変換など）の実用化にも結び付いた。第3次AIブーム（2000年代から現在）では、インターネットやコンピューティングパワーの拡大を背景として、ビッグデータ化と機械学習の進化がブームを牽引し、画像認識・音声認識、機械翻訳、囲碁・将棋などでは人間に追いつき/上回る性能を示し、さまざまなAI応用システム（認識・検索・対話システムなど）が実用化され、社会に普及している。

このような技術発展を図2-1-1では三つの大きな流れで捉えている。

一つ目の流れは「A. 理論の革新」である（図中の紫ライン）。3回のAIブームはいずれも理論面の発展（知識表現・記号処理、辞書・ルールベース処理、機械学習・深層学習など）やコンピューティングパワーの増大などの技術進化によってドライブされた。

二つ目の流れは「B. 応用の革新」である（図中の青ライン）。第2次AIブーム以降は実用的な応用が生まれ始め、ビッグデータの高速並列処理・知識処理の実用化が進み、第3次AIブームでは、機械学習の応用分野が爆発的に拡大した。

三つ目の流れは「C. 社会との関係」である（図中の緑ライン）。これは第3次AIブームを迎えて、活発に議論されるようになった視点である。AI技術のさまざまな応用が社会に広がったことに加えて、AI技術の可能性が人間にとって恩恵だけでなく脅威や弊害ももたらし得るという懸念が強まったためである。

【研究開発の二つの潮流と注目する研究開発領域】

上で述べたような技術発展を経た現在、「第4世代AI」と「信頼されるAI」に向けた取り組みが研究開発の新たな潮流になっている。

現在のAI技術（ここでは「第3世代AI」と呼ぶ）は、さまざまな特定用途において人間を上回る性能を示しているが、大量の学習データ・計算資源が必要であること、学習範囲外の状況に弱いこと、意味処理・説明などの高次処理ができていないこと、といった問題が指摘されている。このような問題の克服に向けて、画像・映像認識や運動制御のような『知覚・運動系のAI技術』と、自然言語処理のような『言語・知識系のAI技術』の融合による「第4世代AI」の研究開発が進み始めた。知覚・運動系、言語・知識系のそれぞれのAI技術においても、深層学習・深層強化学習・深層生成モデル・自己教師あり学習などの技術発展が進んでいることに加えて、『計算脳科学』や『認知発達ロボティクス』の研究から得られる人間の知能に関する知見が「第4世代AI」の研究開発では重要な役割を果たす。そのようなAIと人間あるいは複数のAI間の関係が『エージェント技術』によって広がりを見せている。

その一方で、AI技術が社会に広がり、『社会におけるAI』という視点から、安全性・信頼性・公平性・解釈性・透明性などを含むAI社会原則・AI倫理指針が国・世界レベルで策定され、「信頼されるAI」のための技術開発も重要な研究課題となっている。具体的には、上記原則・指針を満たすようなAI応用システムを開発するための『AIソフトウェア工学』、人がAIと協働してより良い判断や目的達成を目指す『人・AI協働と意思決定支援』、AI・ビッグデータ技術を活用した社会・産業・科学の変革に関わる『AI・データ駆動型問題解決』への取り組みが進展している。

以上において二重力ギカッコ『』で囲った九つを、二つの潮流の中で特に注目する研究開発領域と定める。

2.1

俯瞰区分と研究開発領域 人工知能・ビッグデータ

「第4世代AI」に関わる5領域は、これまでの取り組みからさらなる発展が見られる研究開発領域である。一方、「信頼されるAI」に関わる4領域は、AIの社会への関わりの中で新たに広がってきた研究開発領域である。

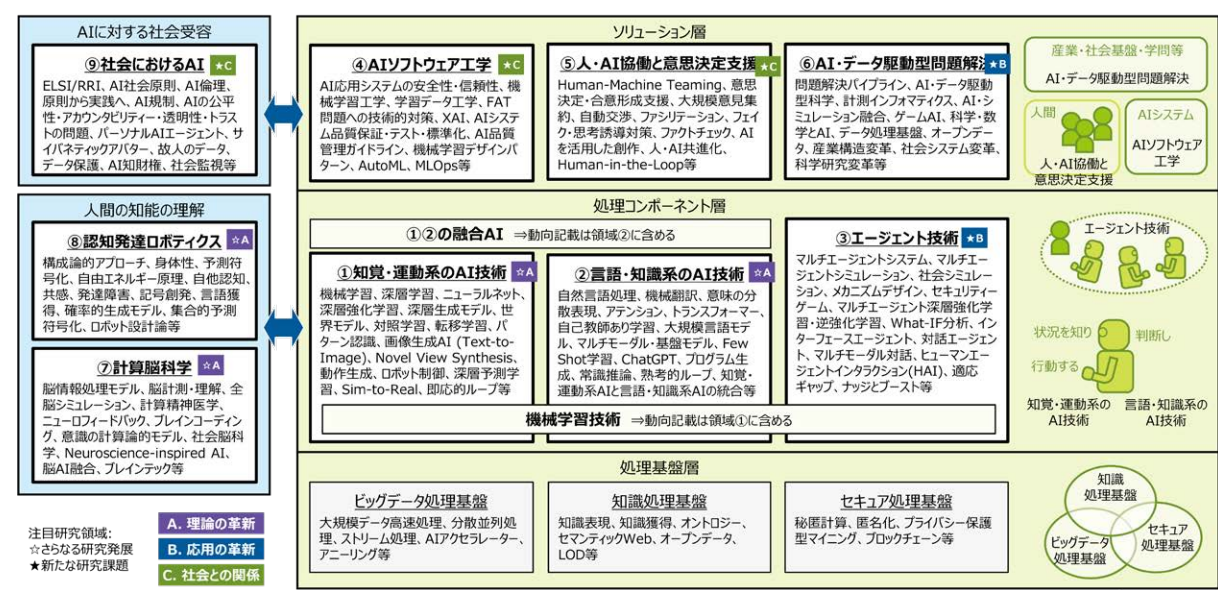


図 2-1-2 人工知能・ビッグデータ技術の俯瞰図（構造）

【AI・ビッグデータの俯瞰図（構造）】

これら九つの注目する研究開発領域を、AI・ビッグデータの技術スタックの中に位置付けた俯瞰図（構造）を図2-1-2に示す。この図では、システムを構成する技術群を「処理基盤層」「処理コンポーネント層」「ソリューション層」に分けている。また、システムを設計する上で、その利用者となる人間やそれが組み入れられる社会についての理解・モデル化および指針の議論も必要になることから、処理コンポーネント層に対応させて「人間の知能の理解」、ソリューション層に対応させて「AIに対する社会受容」という視点を含めた。

図2-1-2の注目する九つの研究開発領域について、その簡単な定義を以下に示す。それぞれの詳細説明は後続の節にまとめている。なお、処理基盤層の技術群は、他の区分（コンピューティングアーキテクチャー区分、セキュリティー・トラスト区分など）で取り上げられるため、ここでは説明対象から外している。

- ①知覚・運動系のAI技術：画像・映像認識に代表される実世界からの入力としての知覚系と、ロボットなどの動作制御に代表される実世界への出力としての運動系という、知能の実世界接点の役割を実現するAI技術に関する研究開発領域である。
- ②言語・知識系のAI技術：自然言語の解析・変換・生成などや知識の抽出・構造化・活用などを行う言語・知識系のAI技術を実現するとともに、知覚系（見る）→言語・知識系（考える）→運動系（動かす）という一連の処理を総合的・統一的な仕組みで実現するための研究開発領域である。
- ③エージェント技術：自ら判断し行動する主体としてのAI（エージェント）について、その自律的メカニズム（自律エージェント）、複数主体の協調（マルチエージェントシステム）、人間とのインタラクション（インターフェースエージェント）、社会的活動・現象のシミュレーション（マルチエージェントシミュレーション）などを実現しようとする研究開発領域である。
- ④AIソフトウェア工学：AI応用システムを、その安全性・信頼性を確保しながら効率よく開発するための新世代のソフトウェア工学を意味する。従来の演繹型システム開発に加えて、機械学習を用いた帰納型システム開発にも対応した開発方法論・技術体系の確立を目指した研究開発領域である。

- ⑤人・AI協働と意思決定支援：人がAIと協働してより良い判断や目的達成を目指す研究開発領域である。特に、個人や集団がある目標を達成するために、考えられる複数の選択肢の中から一つを選択する意思決定を支援するための研究開発を中心に取り上げる。
- ⑥AI・データ駆動型問題解決：AI・ビッグデータ解析が可能にする大規模複雑タスクの自動実行や膨大な選択肢の網羅的検証などによる、問題解決手段の質的变化、産業構造・社会システム・科学研究などの変革を生み出すための研究開発領域である。
- ⑦計算脳科学：脳を情報処理システムとして捉えて、脳の機能を調べる研究開発領域である。人間の知能の情報処理メカニズムの解明、脳疾患・精神疾患の解明や治療、AI技術発展につながる示唆などが期待できる。
- ⑧認知発達ロボティクス：ロボットや計算モデルによるシミュレーションを駆使して、人間の認知発達過程の構成的な理解と、その理解に基づく人間と共生するロボットの設計論の確立を目指した研究開発領域である。
- ⑨社会におけるAI：AI技術が社会に実装されていったときに起こり得る、社会・人間への影響や倫理的・法的・社会的課題 (Ethical, Legal and Social Issues: ELSI) を見通し、あるべき姿や解決策の要件・目標を検討し、それを実現するための制度設計および技術開発を行うための研究開発領域である。

ここで機械学習を、注目する研究開発領域の一つに挙げていないことについて補足しておきたい。機械学習は、現在のAIにおける中核技術であり、上記九つの研究開発領域のほとんどすべてに関わっている。技術スタックの層としては、処理コンポーネント層に相当するので、図2-1-2では、①②③に共通する技術として、機械学習を置いた。そして、後続節においては、機械学習の全般的な動向を「2.1.1 知覚・運動系のAI技術」に記載した上で、他の節でも、その研究開発領域に関わる機械学習のトピックを取り上げることにした。

機械学習を中心に各研究開発領域の動向を見ると、特に深層学習の技術発展が著しく、九つの研究開発領域のそれぞれの発展に大きく影響を与えている。深層学習はまず画像認識・音声認識などのパターン認識に著しい精度改善をもたらした。さらに、強化学習と結び付いた深層強化学習は、試行の繰り返しからアクション決定方策を学習でき、囲碁などのゲームやロボット制御へと応用を広げた。これらは「①知覚・運動系のAI技術」の発展を牽引したが、さらに、意味の分散表現、アテンション、トランスフォーマー、自己教師あり学習といった技術が導入され、深層学習ベースの自然言語処理が大きく進展したとともに、それによって「②言語・知識系のAI技術」と「①知覚・運動系のAI技術」の融合が進み始めたことも注目すべき点である。また、深層生成モデルや自己教師あり学習の発展が①②とも結び付いて、汎用性とマルチモーダル性を高め、画像・映像・音声・文章を自動生成するAI技術が新たな応用の可能性を広げた反面、フェイク生成をはじめAIに対する新たな懸念も引き起こしている。このようなAI・深層学習がもたらす可能性や懸念は、他の研究開発領域③④⑤⑥⑨においても、さまざまなポジティブ/ネガティブな影響を与えている。また、深層学習・深層強化学習・アテンションなどの技術は、人間の知能のメカニズムに通じるものであり、「⑦計算脳科学」「⑧認知発達ロボティクス」の研究成果が、今後も深層学習やAI技術のさらなる発展につながると期待される。

[研究開発状況・戦略の国際比較]

後続の節では、これら九つの研究開発領域の状況を詳しく説明するとともに、領域ごとの国際比較も示している。表2-1-1にその国際比較の部分を抜粋して示した。また、スタンフォード大学が公表しているAI Index¹には、論文・特許・ソフトウェアなどのさまざまな視点から定量的な国際比較が示されている。「1.2.5 研究開発投資や論文、コミュニティ等の動向」にも論文数の国際比較・推移などを示した。これらに見ら

1 <https://aiindex.stanford.edu/report/> (accessed 2023-02-01)

れるように、米国が基礎研究と応用研究・開発の両面で圧倒的優位であるが、中国が急速に追い上げ、研究論文総数では中国が米国を抜くなど、米中2強と言われる状況になっている。

表 2-1-1 AI 研究開発状況の国際比較

国・地域	日本		米国		欧州		中国		韓国	
	基礎	応用	基礎	応用	基礎	応用	基礎	応用	基礎	応用
フェーズ										
①知覚・運動系のAI技術	○↗	○↗	◎↗	◎↗	○→	○↗	○↗	◎↗	△→	△↗
②言語・知識系のAI技術	○↗	○→	◎→	◎↗	○→	○→	◎↗	◎↗	△→	○↗
③エージェント技術	○→	○↗	◎→	◎→	◎→	○→	○↗	○↗	△→	△→
④AIソフトウェア工学	○↗	○↗	○↗	○↗	○↗	○↗	○↗	○↗	×→	×→
⑤人・AI協働と意思決定支援	○↗	○↗	◎↗	◎↗	◎↗	◎↗	◎↗	△→	△→	△→
⑥AI・データ駆動型問題解決	○↗	○↗	◎↗	◎↗	○↗	○→	○→	◎↗	△→	○→
⑦計算脳科学	◎→	○→	◎→	◎→	◎→	◎→	◎↗	○↗	○→	○→
⑧認知発達ロボティクス	○↗	○↗	△→	△↘	○→	○→	△↘	△↘	△↘	△↘
⑨社会におけるAI	◎→	○↗	◎→	◎→	◎→	◎→	△→	△↗	△→	△→

[注] 研究開発領域ごとの状況を相対比較した結果（詳細は後続の各節に記載）を並べたものであり、ある国・地域について研究開発領域間の状況を比較・集計するものではない。

表 2-1-2 AI 技術開発の国際競争状況

	競争戦略のポイント	「第4世代AI」状況	「信頼されるAI」状況
米国	- Big Tech企業がビジネスと基礎研究の両面で圧倒的に優位 - Big Techやスタートアップによる民間の活発な技術開発の一方、DARPAが国としての中長期的な研究投資(AI Next Campaign)をシャープに打ち出している - 経済・国家安全保障のためのAI強化	○革新技術創出から産業化まで強み保有・牽引	○幅広い観点から研究の取り組みがあり、層が厚い
中国	- 次世代人工知能発展計画(AI2030)を掲げ、AIリード企業5社を選定、政府がAI産業を後押し - BtoC中心にビッグデータ獲得、AI実装スピードに勢い - 国際学会でも躍進著しく、米中2強の状況 - 政府はAIを活用した監視・管理社会(社会信用システム、天網、金盾)の構築推進、他国と大きく異なるAI応用技術開発を推進	○第3世代AI技術の改良・実装速度で凌駕	△国として原則は掲げたものの、市場での実践は伴っていない
欧州	- 各国のAI戦略に加えて、Horizon 2020/Europeによる国横断のAI研究(AI for Europe)を推進 - AIに関わる国際ルール作りを通して米中・GAFAと異なる路線を打ち出し(GDPR、信頼できるAIのための倫理ガイドライン、欧州AI法案等)	△強い部分はあるが、米中ほど産業化の勢いが弱い	○理念・倫理ガイドラインを重視・施策化
日本	「人間中心のAI社会原則」「AI戦略2019」(2021年と2022年に更新)を策定、信頼される高品質なAI (Trusted Quality AI)を打ち出し、および、信頼性のある自由なデータ流通に向けたDFFT (Data Free Flow with Trust)を打ち出し - 理研AIP、産総研AIRC、NICTが中核国研として国のAI研究を牽引 日本が優位性を打ち出し得るポイント ✓ Big Techが圧倒しているサイバー世界のAIに対して、日本は製造業・健康医療・モビリティ等を含む実世界AIを重視 ✓ 「Trusted Quality AI」を掲げ、高信頼・高品質な「信頼されるAI」は日本の強みになり得る ✓ AI・ロボット技術に対するマインドの違いを活かす(米国は道具・機械として対峙し、日本は人に寄り添うものととらえる) ✓ 脳科学・発達ロボティクス等の知能の基本的メカニズムの基礎研究、融合AIへの比較的早い取り組み等をもとに「第4世代AI」で先行チャンス	△強い部分はあるが、米中に比べて層が薄い	○信頼性・品質確保のための具体的取り組みでやや先行

AI 技術開発は、産業競争力はもちろん、国の安全保障や社会基盤をも支えるものと認識されるようになり、各国ともAI 技術開発の強化戦略を打ち出し、産業界での技術開発推進と国による戦略的研究投資の強化が図られている。この国際競争状況を、米国・中国・欧州・日本の比較という形で表 2-1-2 にまとめた。ここではこの表の詳細な説明は省くが（後続節の中で触れている）、表 2-1-1 に示した研究開発状況や競争優位分野の差異は、表 2-1-2 に示したような競争戦略が背景にある。

また、表 2-1-2 には、「第 4 世代 AI」と「信頼される AI」という二つの潮流の観点からの状況比較も示した。

日本が国際競争力を確保するためにも、この二つへの取り組み方や推進策が重要である。「信頼されるAI」への取り組みについては、「AI戦略2019」で「信頼される高品質なAI」(Trusted Quality AI) が打ち出され、具体的な取り組みも国際的にやや先行している。品質へのこだわりは日本の産業界が伝統的に持っている強みでもあり、AI関連産業においても競争力になると期待される。「第4世代AI」への取り組みについては、深層学習を中心とする現在のAI(第3世代AI)が米中2強の状況であっても、競争の土俵が変わるとすれば、日本にも先行するチャンスがある。

【国としての研究開発強化の方向性】

上で述べたように、「第4世代AI」と「信頼されるAI」への取り組みは、研究開発の二つの潮流と日本の国際競争力の強化という観点から特に重要である。

第1の潮流「第4世代AI」は、第3世代AIの中核である深層学習の抱える問題点(大規模なデータと計算資源が必要なこと、想定外の状況に臨機応変に対応できないこと、説明や意味など高次の処理ができていないこと)を克服する新しいAIのアーキテクチャーを探究する方向性である。大規模学習によって、マルチモーダル性と汎用性を大きく向上させた基盤モデル(Foundation Model)が登場し、大きく注目されている。しかし、この力任せのアプローチは人間の知能とは異なる方向に向かうように思える。人間は大量の教師データがなくとも発達・成長するし、学習した結果を状況に応じて臨機応変に応用できる。人間の知能から学ぶアプローチとして、即応的な知能(システム1)と熟考的な知能(システム2)から成る二重過程モデルや、身体性や環境とのインタラクションを通した予測誤差最小化原理に基づく発達・創発モデルが注目され、研究が活発化しつつある。このような基盤モデル、二重過程モデル、発達・創発モデルを中心とした「知能モデルの解明・探究」が、「第4世代AI」に向けた重要な研究開発課題である。

第2の潮流「信頼されるAI」については、AIに対する社会からの要請が原則から実践フェーズへ移行する中で、技術開発だけでは必ずしも要請に十分に答えられない問題が顕在化してきた。AIの信頼性だけでなく、社会におけるトラスト(信頼)を、対象真正性、内容真実性、振る舞い予想・対応可能性といった多面から捉えて複合的に取り組むことが必要である。そのため、技術開発だけでなく制度設計や人文・社会科学の知見も結集した総合知による「デジタル社会におけるトラスト形成」が、今後重要な研究開発課題になる。

上述の「第4世代AI」「信頼されるAI」という二つの潮流と、この1-2年の基盤モデル・生成AIの急激な発展を踏まえて、JST CRDSでは2024年3月に、さらなるAIの発展方向と基礎研究課題について調査・検討を行い、「次世代AIモデルの研究開発」に関する戦略プロポーザルを公開した²。この戦略プロポーザルでは、現在の基盤モデル・生成AIは、資源効率、実世界操作(身体性)、論理性・正確性、安全性・信頼性に課題があることを指摘し、それらを克服するための研究開発として、幅広に複数通りのアプローチを進め、それらの間の融合やシナジーを生み出す取り組みの必要性を述べている。また、さまざまな生成AIリスクを指摘し、そういったAIリスクに対して五つのケースで対策を創り出していく必要性を述べている。また、AIに対する社会からの要請、および、AIの発展による社会変革の両面を考えていく必要性から、目指したい社会の在り方に関する哲学を含む人文・社会科学面からの研究・知見や、AIによる社会・産業・科学のプロセス変革への取り組みなどについても検討している。

また、「AI戦略2019」においても中核的研究開発課題に設定された「信頼されるAI」が依然重要であることに加えて、「AI戦略2022」では差し迫った危機への対処や社会実装の推進などAI活用が強調されている。上述の「第4世代AI」「信頼されるAI」を含む次世代AIモデルの研究開発成果が、社会のさまざまな場面で活用されるためには、社会スケールの計算アーキテクチャー、それを支えるデータ基盤・データエコシステム、AIエージェント分散協調メカニズムなどを含む「社会システムを支えるAIアーキテクチャー」が、もう一つ

2 <https://www.jst.go.jp/crds/report/CRDS-FY2023-SP-03.html>

の重要な研究開発課題になる。このような視座に立つならば、次世代AIモデルとして検討したようなAI自体の技術発展から、むしろ、人やAI（エージェント）が多数混在し、相互作用しながら共生・協働する「人・AI共生社会」への発展を描くことが必要である。そのための社会基盤の要件やそれを実現するための技術開発を考えていくことが、次の重要課題になると考える。

2.1

俯瞰区分と研究開発領域
人工知能・ビッグデータ