科学技術振興機構
Japan Science and Technology Agency

CRDS 研究開発戦略センター
Center for Research and Development Strategy

参考資料7
科学技術・学術審議会
情報委員会（第47回）
令和8年5月12日

俯瞰報告書2026 システム・情報科学技術分野について

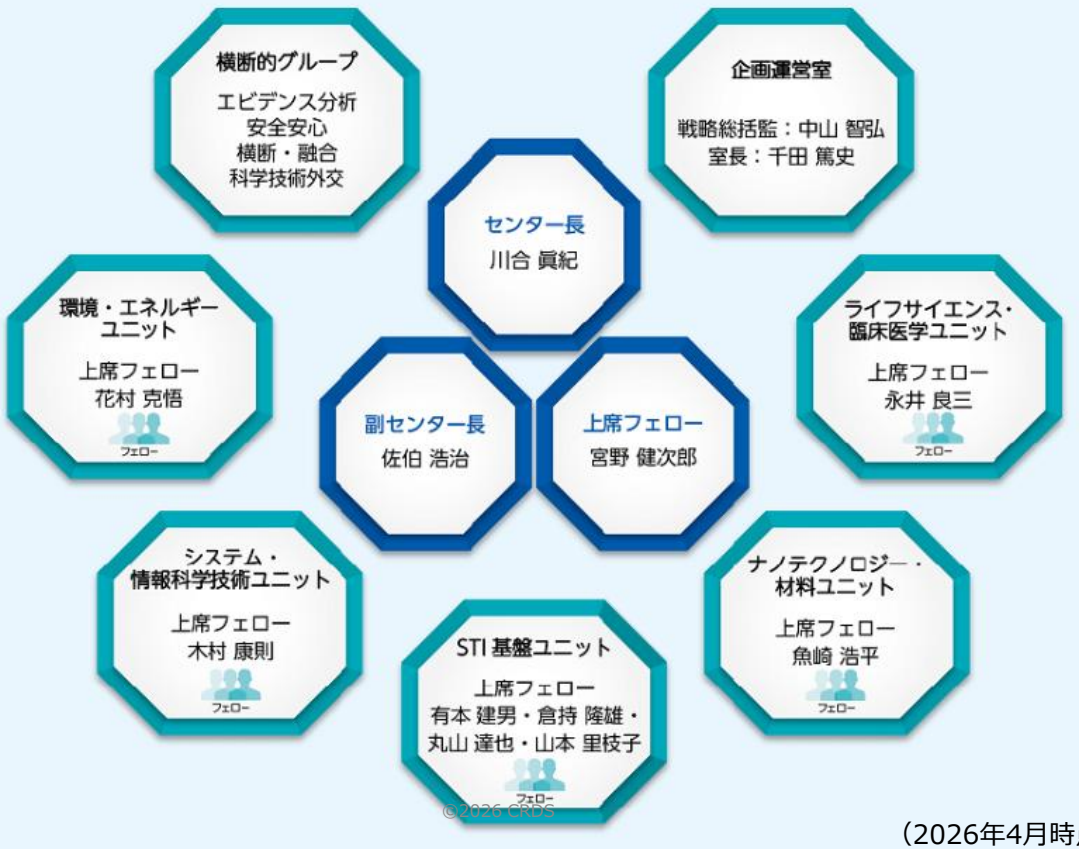
2026年5月12日

JST研究開発戦略センター
システム・情報科学技術ユニット

研究開発戦略センター（CRDS）について

- 国の科学技術イノベーション政策に関する調査、分析、提案を中立的な立場に立って行う組織として、平成15年（2003年）7月設置
- 国内外の科学技術イノベーションや関連する社会および政策動向を俯瞰的に調査・分析し、課題の抽出を行い、各種の科学技術イノベーション政策や研究開発戦略に資する情報を提案し、その実現に向け活動している公的シンクタンク

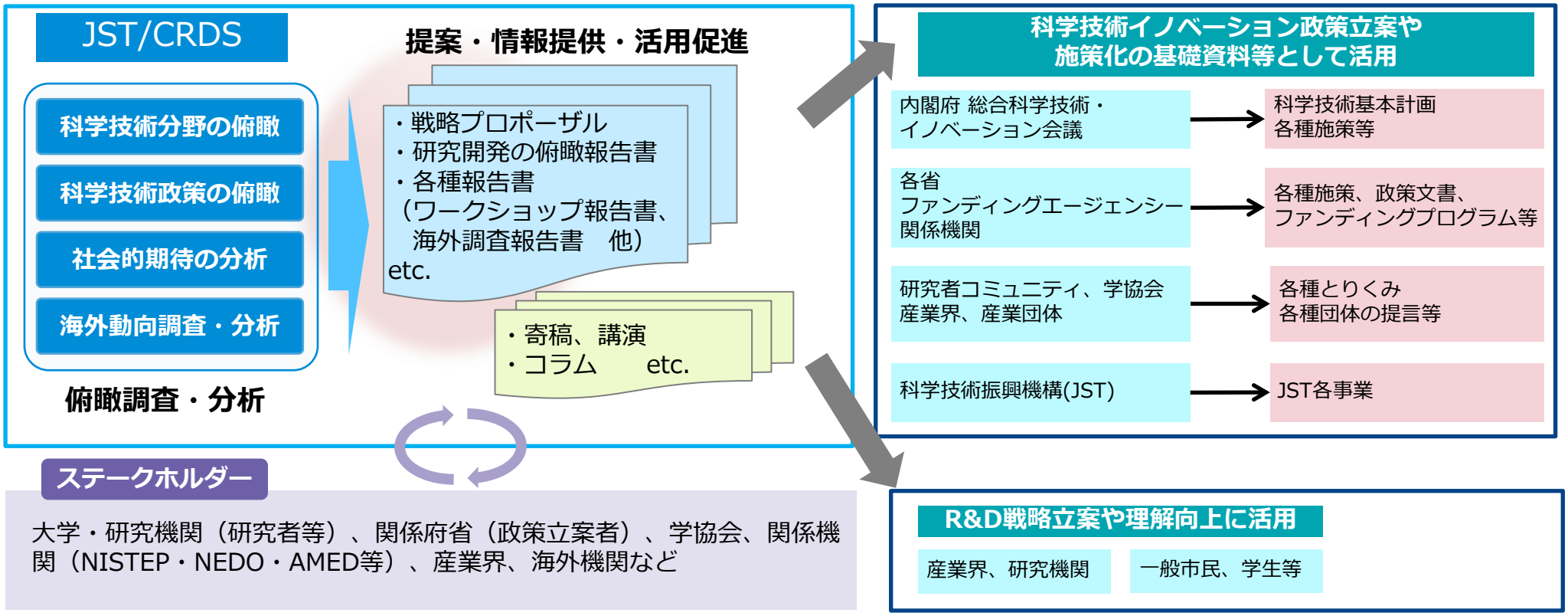
推進体制 ～総合知の創出・活用を目指して～



（2026年4月時点）

研究開発戦略センター（CRDS）活動プロセス

①国内外の社会や科学技術イノベーションの動向及びそれらに関する政策動向を俯瞰し、分析しています。
②「俯瞰報告書」や研究開発戦略提案「戦略プロポーザル」をとりまとめ、提案の実現に向けた取組を行っています。
③ワークショップ等を開催し、関係者の共通認識の醸成を図っています。



研究開発の俯瞰報告書 分野別俯瞰編（2026年）概要

システム情報科学技術分野

研究開発や経済活動から日常生活に至るまで、AIを含むシステムや情報利用が深く浸透。多様な分野の研究開発のみならず、人類の営み全般に変革をもたらしつつある一方で、新たなリスクも顕在化。AIに必要な膨大な計算資源と電力消費に応える情報基盤が課題に。

（世界の研究開発動向）

■ICT分野の研究開発は「AIの自律化・フィジカル化・汎用化」と「AIトランスフォーメーションと多層化するリスクへの対応」の潮流を「知能革命を支えるAIインフラ」の潮流が支える。

- AIの自律化・フィジカル化・汎用化：AIは、人間の指示に従うツールから、自ら環境を認識し、判断し、行動する自律的な動作可能へと進化を遂げつつある。また、ロボットやドローンに実装されることで、フィジカルな存在として社会の中に溶け込んでいく。さらに、従来のタスク特化型から脱却し、学習・推論・創造を横断的にこなす汎用的な知能の萌芽が生成AIにおいて見え始めている。
- AIトランスフォーメーションと多層化するリスクへの対応：AIは単なる効率化の手段ではなく、研究・教育・創作・産業活動といった人間の知的活動そのものを変革する存在になりつつある。科学の進め方やビジネスの構造、人間の学び方にまで抜本的な変革をもたらす。しかし同時に、AIが社会のあらゆる層に浸透することは、新たなリスクを多層的に顕在化させる。
- 知能革命を支えるAIインフラ：知能革命を持続的に推進するには、AIの特性に即した新しい情報基盤が必要である。膨大な計算資源と電力を消費する従来型の手法に依存するのではなく、確率的推論や近似解を活用するAIの特性を最大限に引き出す省エネルギー型アーキテクチャーや、エッジ・クラウド・分散コンピューティングを組み合わせた柔軟な計算インフラが求められる。

（主要国の政策動向）

日本	• 統合イノベーション戦略2025発表。AIイノベーション促進とリスク対応の両立、次世代情報通信基盤の開発・導入、量子技術研究開発の推進を記載。 • 人工知能関連技術の研究開発及び活用の推進に関する法律（通称：AI新法）公布。AI技術の研究開発と社会実装を適切に推進。
米国	• 米国AI行動計画を発表（2025年7月）。「安全性・信頼性重視」のバイデン路線から、「規制排除・自由な成長優先」への政策シフト。 • 国家量子イニシアチブ法が成立（2018年）。国家量子イニシアチブ諮問委員会（NQIAC）、国家量子調整室（NQCO）などが設定。
欧州	• デジタル市場法（DMA）とデジタルサービス法（DSA）が2022年に発効 • AI法が2024年に発効。AI大陸行動計画を2025年に発表
中国	• 人工知能＋（AIプラス）行動計画を提示（2024年）。デジタル産業クラスターを建設。 • ロボット＋応用行動計画を公布（2023年）

（日本の研究開発動向）

■Society 5.0実現に向けた研究開発
- 多様な活動を支えるサイバネティック・アバター技術と社会基盤の構築を推進。

■AIモデルの研究開発
- 国立情報学研究所（NII）のLLMCを中核機関として、国内産学の研究者・開発者を結集したLLM-jpプロジェクトで大規模言語モデルを構築。
- 計算脳科学とAIの融合研究や、日本発の研究領域である認知発達・記号創発ロボティクス等「次世代AIモデル」につながる基礎研究も推進。

■「安全・安心で信頼できるAI」を掲げ国際標準化活動でも健闘
- 研究コミュニティーや産業界での取り組みを国際的に早い時期から開始。
- 実践的なトラストを積み上げる取り組みを推進。

■スーパーコンピューター「富岳」の成果創出と「富岳NEXT」の開発開始
- 「AI・データ科学との融合・連携」など4領域20課題で新規成果の創出に取り組む。
- 計算による課題解決を支える次世代「AI-HPCプラットフォーム」開発体制が始動。

（わが国として重要な研究開発）

■AI社会のリスク対策・トラスト基盤
- AIの自律性が高まることで生じる多層的なリスクへの対策と、トラストの確保

■コグニティブセキュリティー
- フェイク等の情報攻撃から人と社会を守り、安心安全な社会を支える

■AIトランスフォーメーション基盤
- AIトランスフォーメーションのための共通基盤技術の開発および分野適用

■次世代AIモデル
- 人・AI共生モデル、人間知能理解、資源・エネルギー効率の観点から追求

■フィジカルAIシステム
- 身体性に基づく知能の研究と環境を理解し人間と協調するロボット実現

■バイオハイブリッドロボット
- 生体・生体材料と人工物を組み合わせ生体特有の運動や感覚も持つロボット実現

■因果推論
- 従来の機械学習が回答できない因果関係をデータから抽出する基盤技術の研究

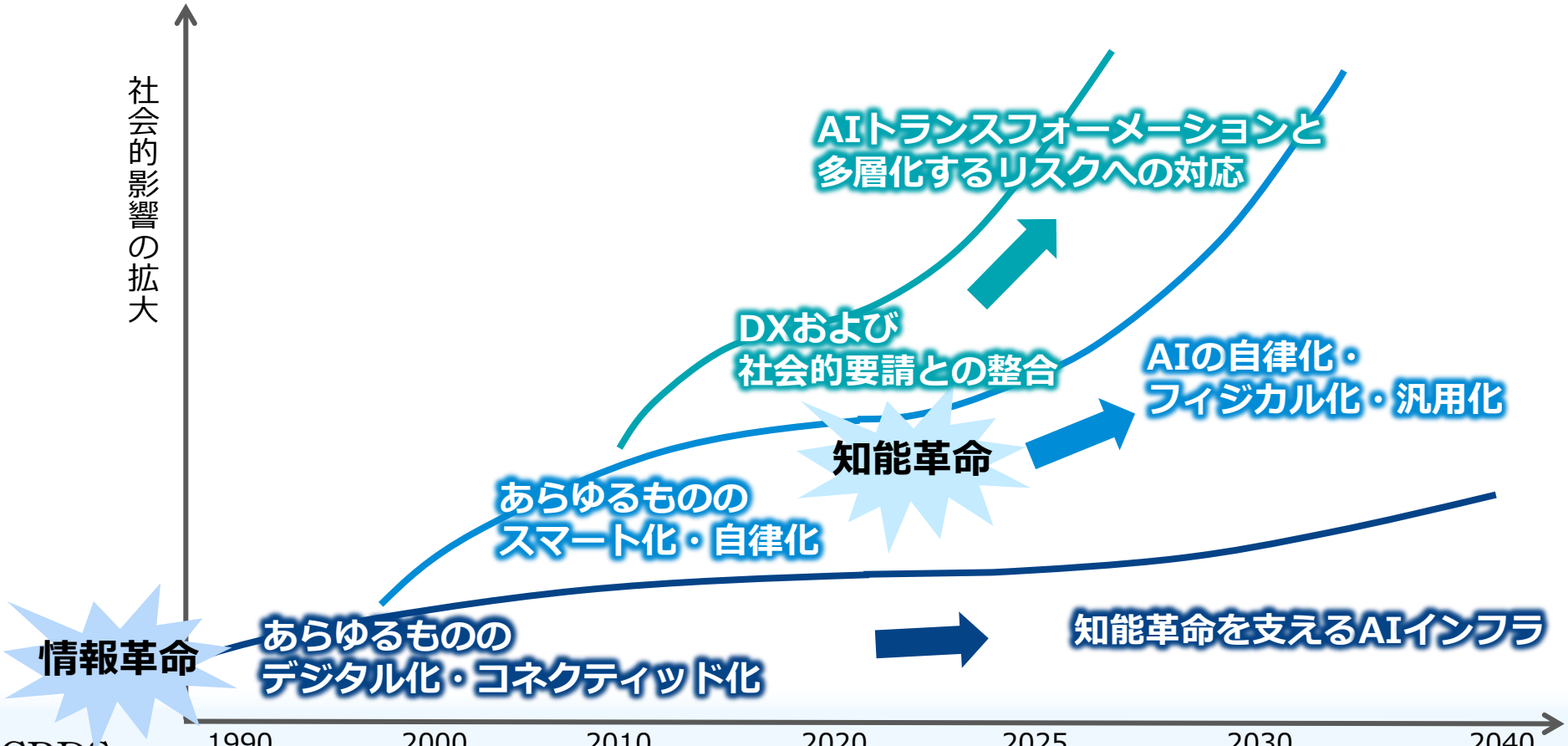
■最適化
- 数理最適化と機械学習の融合による最適化手法の研究開発と社会実装

■通信と計算の融合
- 通信と計算を一体化した将来ネットワーク

■AIコンピューティング
- AI処理を高速かつ省電力で実行する計算基盤

世界の研究開発動向

- 生成AIの登場によって人の知的活動が劇的に変革（**知能革命**）し、ICT全体の潮流がAIを中心とする流れに大きく変化
- 「AIの自律化・フィジカル化・汎用化」と「AIトランスフォーメーションと多層化するリスクへの対応」の2つの潮流を「知能革命を支えるAIインフラ」の潮流が支える



世界の研究開発動向

- **生成AIの登場によって人の知的活動が劇的に変革（知能革命）し、ICT全体の潮流がAIを中心とする流れに大きく変化**
- **「AIの自律化・フィジカル化・汎用化」と「AIトランスフォーメーションと多層化するリスクへの対応」の2つの潮流を「知能革命を支えるAIインフラ」の潮流が支える**

AIトランスフォーメーションと多層化するリスクへの対応

- AIが単なる効率化の手段から、研究・教育・創作・産業活動といった人間の知的活動そのものを変革する存在に。科学の進め方やビジネスの構造、人間の学び方まで抜本的に変革。
- 一方、AIが社会のあらゆる層に浸透することで新たなリスクを多層的に顕在化。

AIの自律化・フィジカル化・汎用化

- 自律化：AIが人間の指示に従うツールから、自ら環境を認識し、判断し、行動する自律的な動作可能へと進化を遂げつつある。
- フィジカル化：AIがロボットやドローンに実装されることで、フィジカルな存在として社会の中に溶け込んでいく。
- 汎用化：従来のタスク特化型から脱却し、学習・推論・創造を横断的にこなす汎用的な知能の萌芽が生成AIにおいて見え始めている。

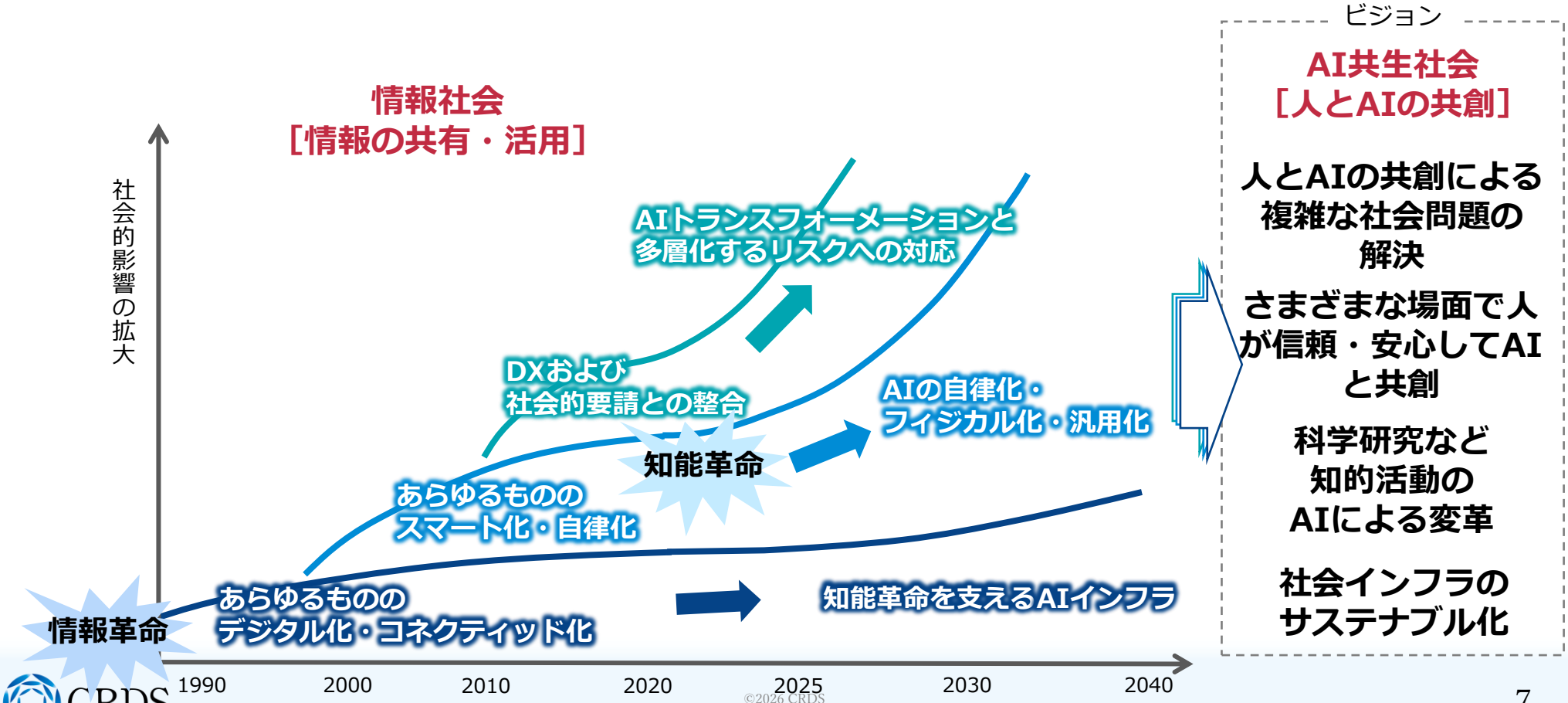
知能革命を支えるAIインフラ

- 知能革命を持続的に推進するには、AIの特性に即した新しい情報基盤が必要。
- 膨大な計算資源と電力を消費する従来型の手法に依存しない計算インフラが求められる。
- AIが扱うデータの膨張に応じて、高速・高効率かつ低環境負荷なネットワークやストレージ技術も不可欠。

ICTが目指すべきビジョン「AI共生社会」

「情報革命」によってもたらされた、人が情報を意識的に共有・活用する「**情報社会**」

「知能革命」によってもたらされる、AIと調和し共に歩む「**AI共生社会**」へ



研究分野俯瞰の方法　対象とする分野・領域の設定

科学を支え、未来へつなぐ

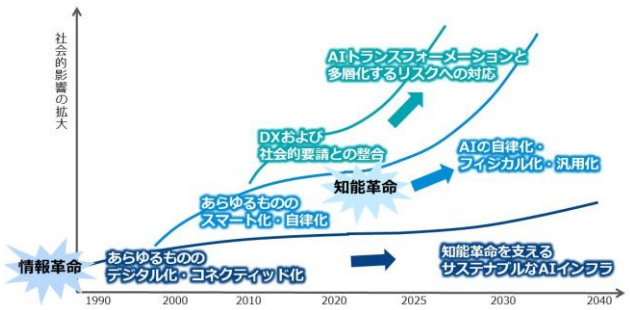
注目技術群の母集合 → **3つの選定基準** → **俯瞰対象とする注目すべき研究領域を決定**

- 外部の調査会社・団体・学会等が注目する技術トピック群
- 有識者ヒアリング等に基づきCRDS関係メンバーが注目した技術トピック群
- CRDS特任フェローによるアドバイス

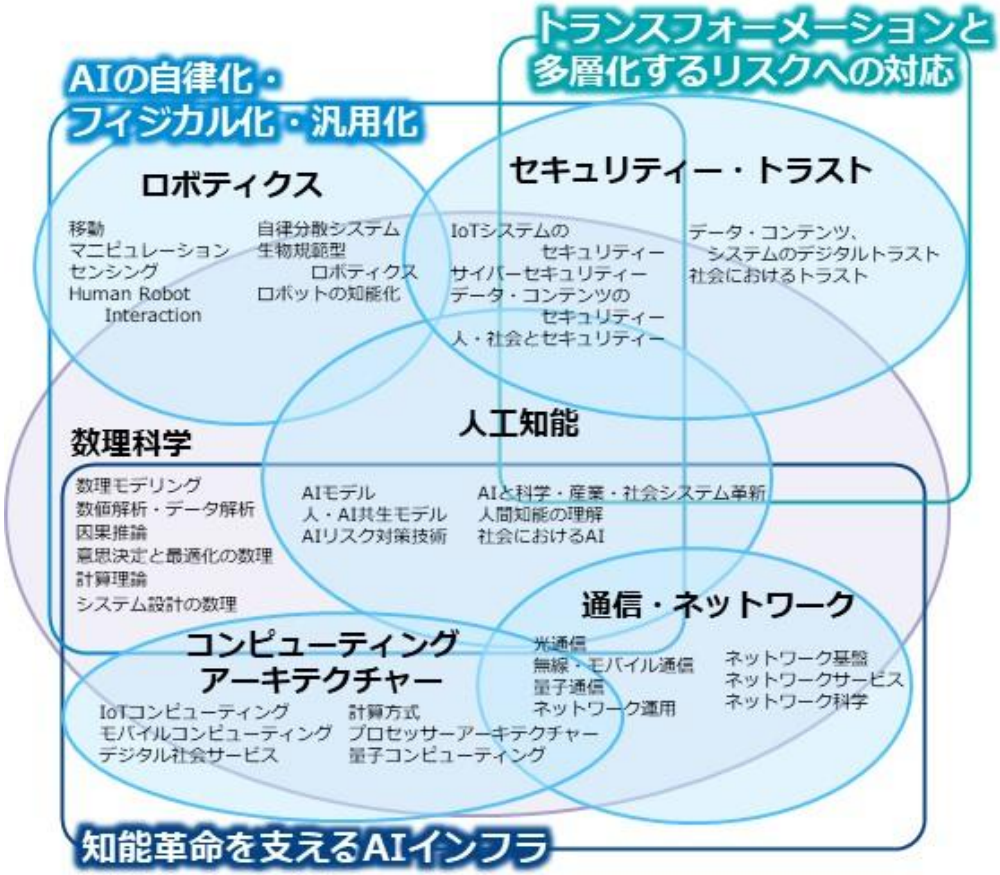
（1）エマージング性
技術の革新性やその技術への期待の急速な高まりに注目

（2）社会の要請・ビジョン
社会からの要請や国のビジョンとの整合性に着目

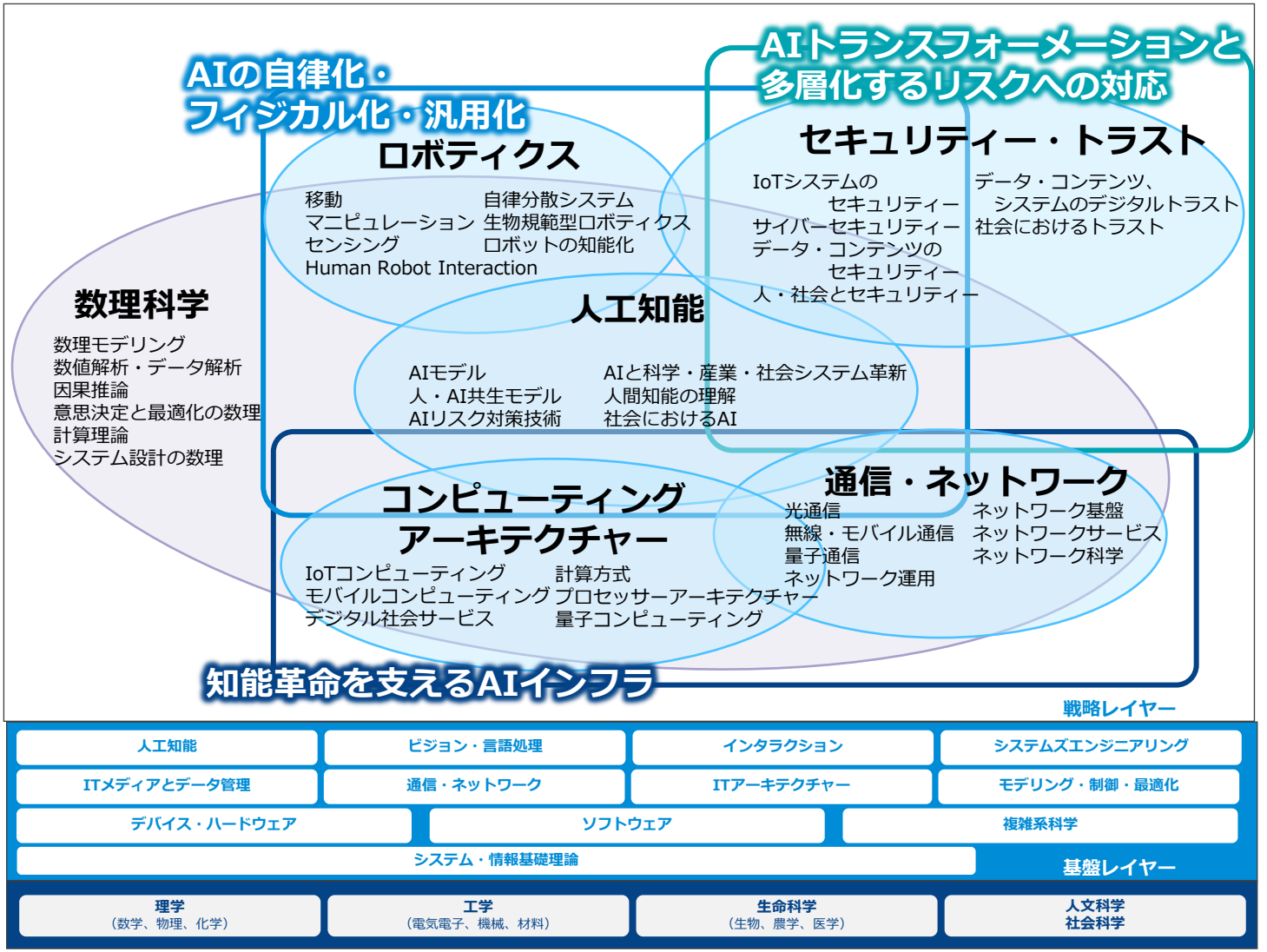
（3）社会インパクト
人々のライフスタイル・ワークスタイルや社会・産業構造の変革、SDGsを含む社会課題解決への貢献に着目



2025年の断面

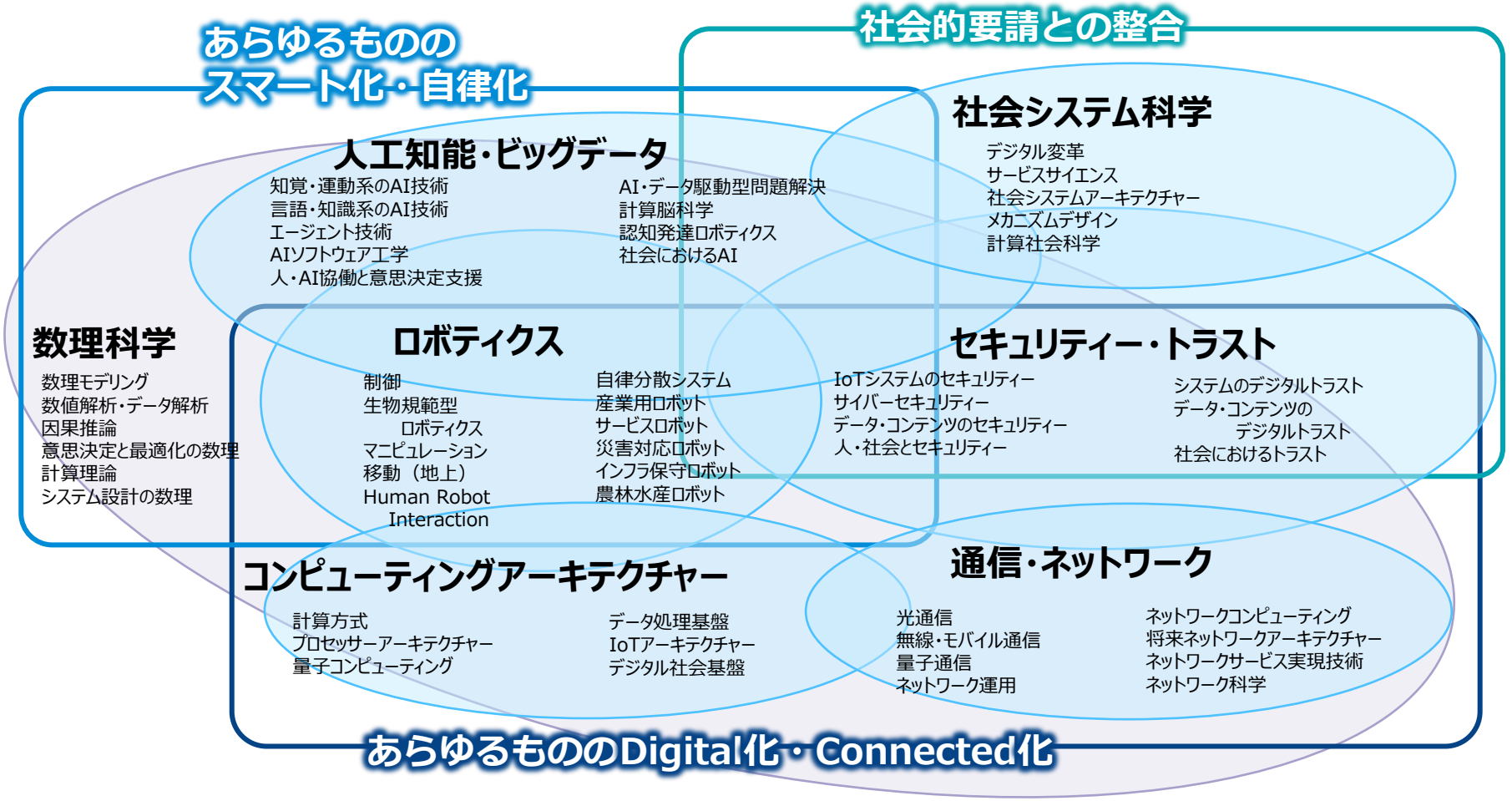


システム・情報科学技術の俯瞰図



- 基盤レイヤー：既に学問分野として確立された区分に基づき、基盤技術として世界に通用するものを生み出すための研究開発に着眼
- 戦略レイヤー：社会の要請・ビジョンと技術のトレンドの両者を鑑み、「エマージング性」「社会の要請・ビジョン」「社会インパクト」の3点を基準として戦略的な重要度が高い研究開発領域を複数特定したもの

【参考】研究開発領域（2024） 7区分53 研究開発領域



俯瞰対象とする区分・研究開発領域（リスト）

俯瞰区分	研究開発領域
人工知能（AI）	人間知能の理解
	AIモデル
	人・AI共生モデル
	社会におけるAI
	AIリスク対策技術
	AIと科学・産業・社会システム革新 (AIトランスフォーメーション)
ロボティクス	制御
	生物規範型ロボティクス
	マニピュレーション
	移動
	Human Robot Interaction
	自律分散システム
	ロボットの知能化
セキュリティー・トラスト	IoTシステムのセキュリティー
	サイバーセキュリティー
	データ・コンテンツのセキュリティー
	人・社会とセキュリティー
	データ・コンテンツ・システムのデジタルトラスト
	社会におけるトラスト

俯瞰区分	研究開発領域
コンピューティングアーキテクチャー	計算方式
	プロセッサーアーキテクチャー
	量子コンピューティング
	IoTコンピューティング
	モバイルコンピューティング
	デジタル社会サービス
通信・ネットワーク	光通信
	無線・モバイル通信
	量子通信
	ネットワーク運用
	ネットワーク基盤
	ネットワークサービス
	ネットワーク科学
数理科学	数理モデリング
	数値解析・データ解析
	因果推論
	意思決定と最適化の数理
	計算理論
	システム設計の数理

主要国の政策動向

日本	• 統合イノベーション戦略2025発表。AIイノベーション促進とリスク対応の両立、次世代情報通信基盤の開発・導入、量子技術研究開発の推進を記載 • 人工知能関連技術の研究開発及び活用の推進に関する法律（通称：AI法）公布。AI技術の研究開発と社会実装を適切に推進
米国	• 米国AI行動計画を発表（2025年7月）。「安全性・信頼性重視」のバイデン路線から、「規制排除・自由な成長優先」への政策シフト • 国家量子イニシアチブ法が成立（2018年）。国家量子イニシアチブ諮問委員会（NQIAC）、国家量子調整室（NQCO）などが設定
欧州	• デジタル市場法（DMA）とデジタルサービス法（DSA）が2022年に発効 • AI法が2024年に発効。AI大陸行動計画を2025年に発表
中国	• 人工知能＋（AIプラス）行動計画を提示（2024年）。デジタル産業クラスターを建設 • ロボット＋応用行動計画を公布（2023年）

日本の研究開発動向

AIの自律化・フィジカル化・汎用化

■AIモデルの研究開発

- 国産のオープンな大規模言語モデル開発のため、国立情報学研究所(NII)のLLMCを中核機関として、国内産学の研究者・開発者を結集したLLM-jpプロジェクトが進行中
- AI×ロボットの技術開発では、AIロボット協会(AIRoA)が発足し、国産のオープンなロボット基盤モデル開発を推進中、日本発の認知発達・記号創発ロボティクスの取り組みも

AIトランスフォーメーション(AX)と多層化するリスクへの対応

■「安全・安心で信頼できるAI」を掲げ国際標準化活動でも健闘

- 研究コミュニティや産業界の取り組みを早期開始、広島AIプロセスなど国際的に先導

■科学のAX「AI for Science」の推進体制を強化

- Nobel Turing Challengeを世界に先駆けて提唱、第7期基本計画で「科学の再興」を推進

知能革命を支えるAIインフラ

■スーパーコンピューター「富岳」の成果創出と「富岳NEXT」の開発開始

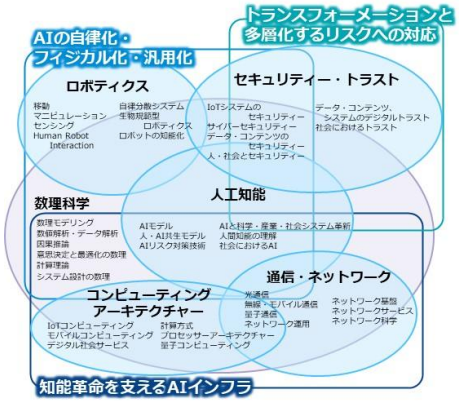
- 「AI・データ科学との融合・連携」など4領域20課題で新規成果の創出に取り組む。
- 計算による課題解決を支える次世代「AI-HPCプラットフォーム」開発体制が始動

わが国として重要な研究開発 俯瞰から戦略提言へのロジック

科学を支え、未来へつなぐ

俯瞰対象とする

注目すべき研究領域



総合的な戦略精査

- **4つのシナリオ**※に基づく国際競争力確保
- **技術トレンド/ビジョン**に沿った強みの継続的育成・蓄積
- **社会環境・世界情勢の変化**に応じた打ち手
- **ユニット外有識者との議論**

国として推進すべき

重点テーマ候補

		①強い技術	②強い産業	③社会受容	④経済安保
AIトランスフォーメーションと多層化するリスクへの対応	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ー			●	●
	AIトランスフォーメーション基盤		●		●
AIの自律化・フィジカル化・汎用化	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	●			●
知能革命を支えるAIインフラ	通信と計算の融合	●			●
	AIコンピューティング		●		●

※国際競争力確保に向けた **4つのシナリオ**

①強い技術を核とした骨太化 既に保有している、あるいは、育ちつつある**強い技術を足掛かりとして**、技術の国際競争力を骨太化する作戦	**③社会課題の先行解決** 課題先進国として、先端技術の**社会受容性で先行できることを活かして**、国際競争力を構築する作戦
②強い産業の発展・革新の推進 既に保有している、あるいは、育ちつつある**強い産業を足掛かりとして**、国際競争力のある技術群を育てる作戦	**④経済安全保障上の重要基盤技術の確保** **経済安全保障上の重要基盤技術については**、国として保有・強化しなくてはならないという考え

わが国として重要な研究開発　重点テーマ概要①

AIの自律化・フィジカル化・汎用化

次世代AIモデル	人間や社会にとってどのようなAIを目指すべきか、人・AI共生モデルの観点、人間知能の理解の観点、資源・エネルギー効率の観点なども含めて、最先端のAIモデルを追求するための研究開発テーマ
フィジカルAIシステム	現在の深層学習の課題を克服し、人間と親和性が高く、実世界で発達・成長する、新しいAIの理論・アーキテクチャーにつながる研究開発テーマ
バイオハイブリッドロボット	生体もしくは生体材料からできた部品と人工物からできた部品を組み合わせて、生体特有の運動や感覚といった機能をアクチュエーターやセンサーとして利用したロボットに関する研究開発テーマ
因果推論	データから因果関係を抽出する基盤技術である因果推論に関する研究開発テーマ
最適化	ひろく社会において最適な行動や手段、設計などを定める指針を与えるための研究開発テーマ

わが国として重要な研究開発　重点テーマ概要②

AIトランスフォーメーションと多層化するリスクへの対応

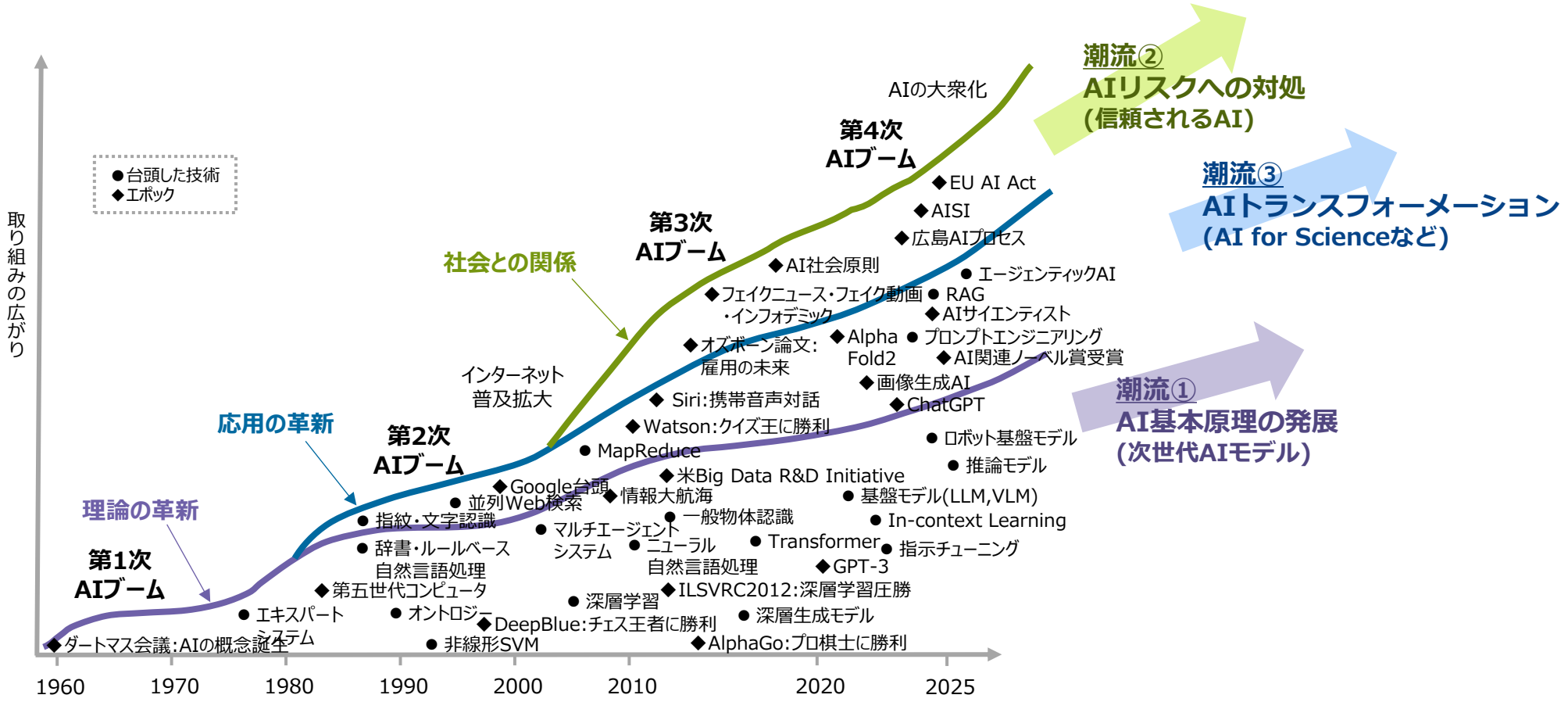
AI社会のリスク対策・トラスト基盤	AIの進展に伴い顕在化しているリスクへの対策と、デジタル化に伴って揺らいでいるトラスト（信頼）の確保を目指す総合的な研究開発テーマ
コグニティブセキュリティー	人間の認知や思考、意思決定などに悪影響を与える攻撃からの防御に関する研究開発テーマ
AIトランスフォーメーション基盤	科学研究におけるAI活用（AI for Science）をはじめとした、さまざまな分野における革新（AIトランスフォーメーション）のためのフレームワークやそれを支える共通基盤技術の開発、および、その分野適用のための研究開発テーマ

知能革命を支えるAIインフラ

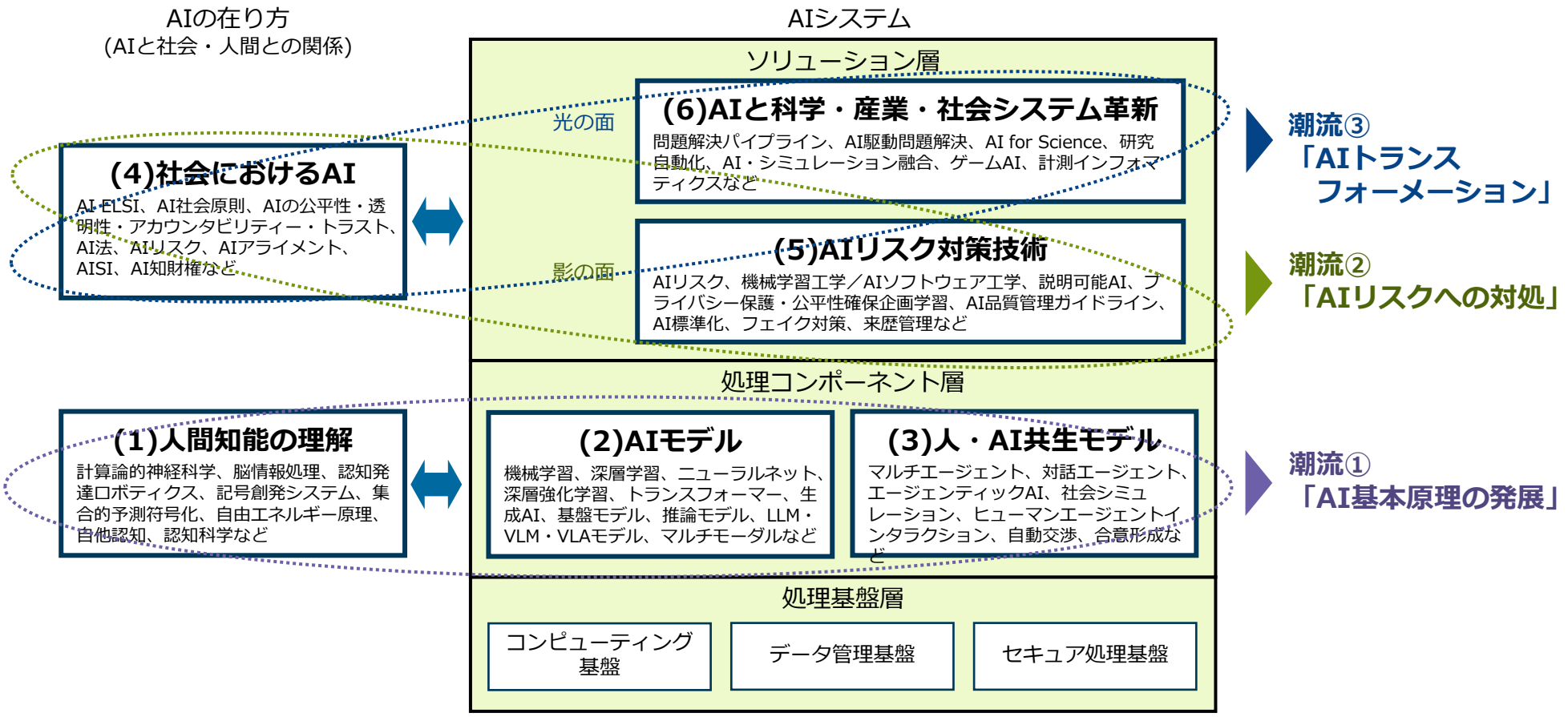
通信と計算の融合	AIを含む計算処理により、通信周波数の割り当てや通信経路の決定を最適化したり、計算（AI、データ処理、アプリケーション・サービス）を効率的に実行できるように通信ネットワークを構成するなど、通信と計算とを一体的に考えて技術確立を目指す研究開発テーマ
AIコンピューティング	AI時代に即した計算方式、コンピューターアーキテクチャー、データセンターおよび広域分散処理技術に関する研究開発テーマ

参考資料①
各区分の時系列・構造俯瞰図

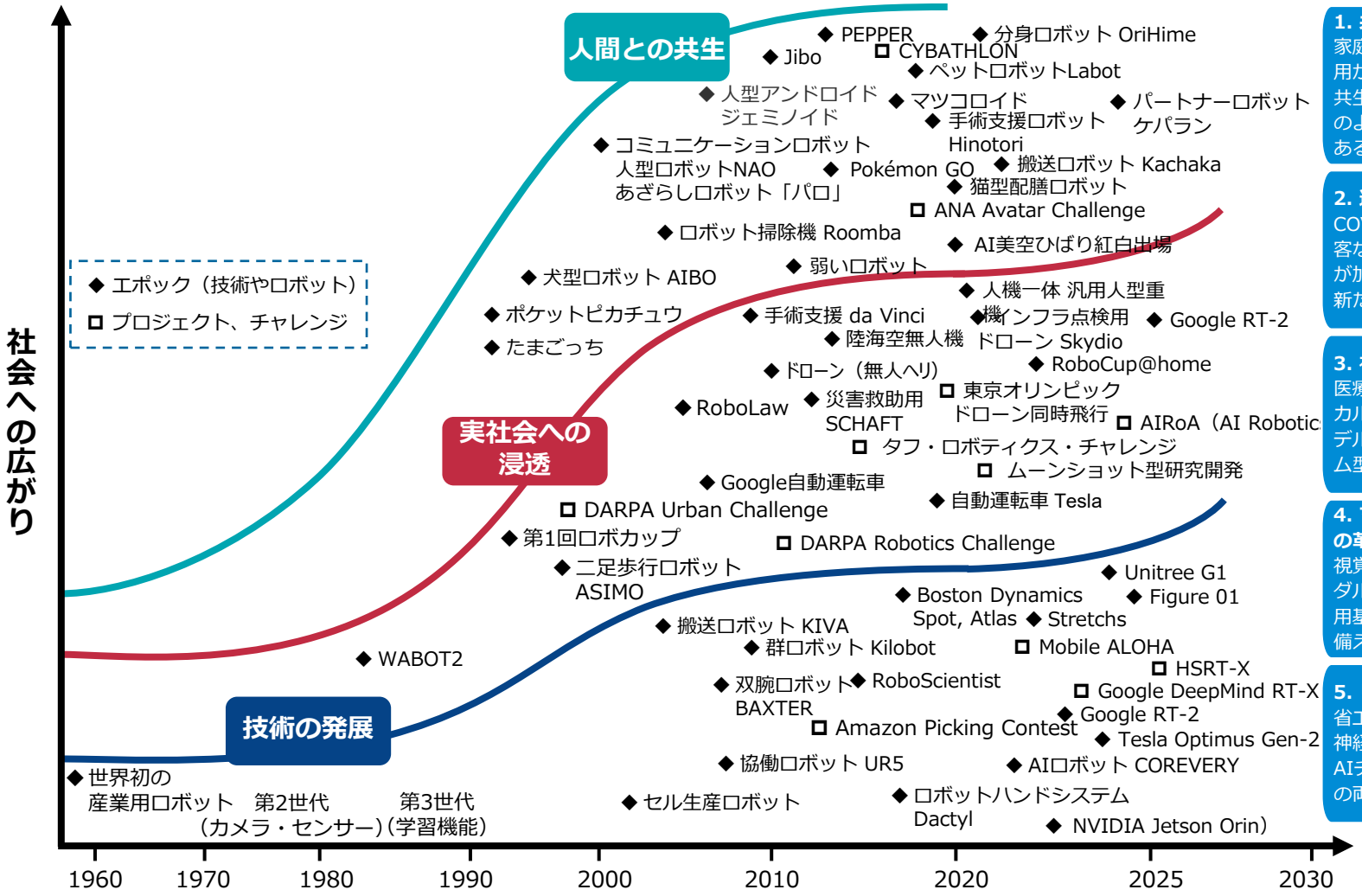
【人工知能区分】時系列俯瞰図



【人工知能区分】構造俯瞰図



【ロボティクス区分】時系列俯瞰図



近年の顕著な動向

1. 共生的インタラクションの社会実装
家庭・公共空間におけるサービスロボットの活用が進展。感情的応答や日常的接点に着目した共生志向の設計が重視され、RoboCup@Homeのような生活環境ベンチマークも整備されつつある。

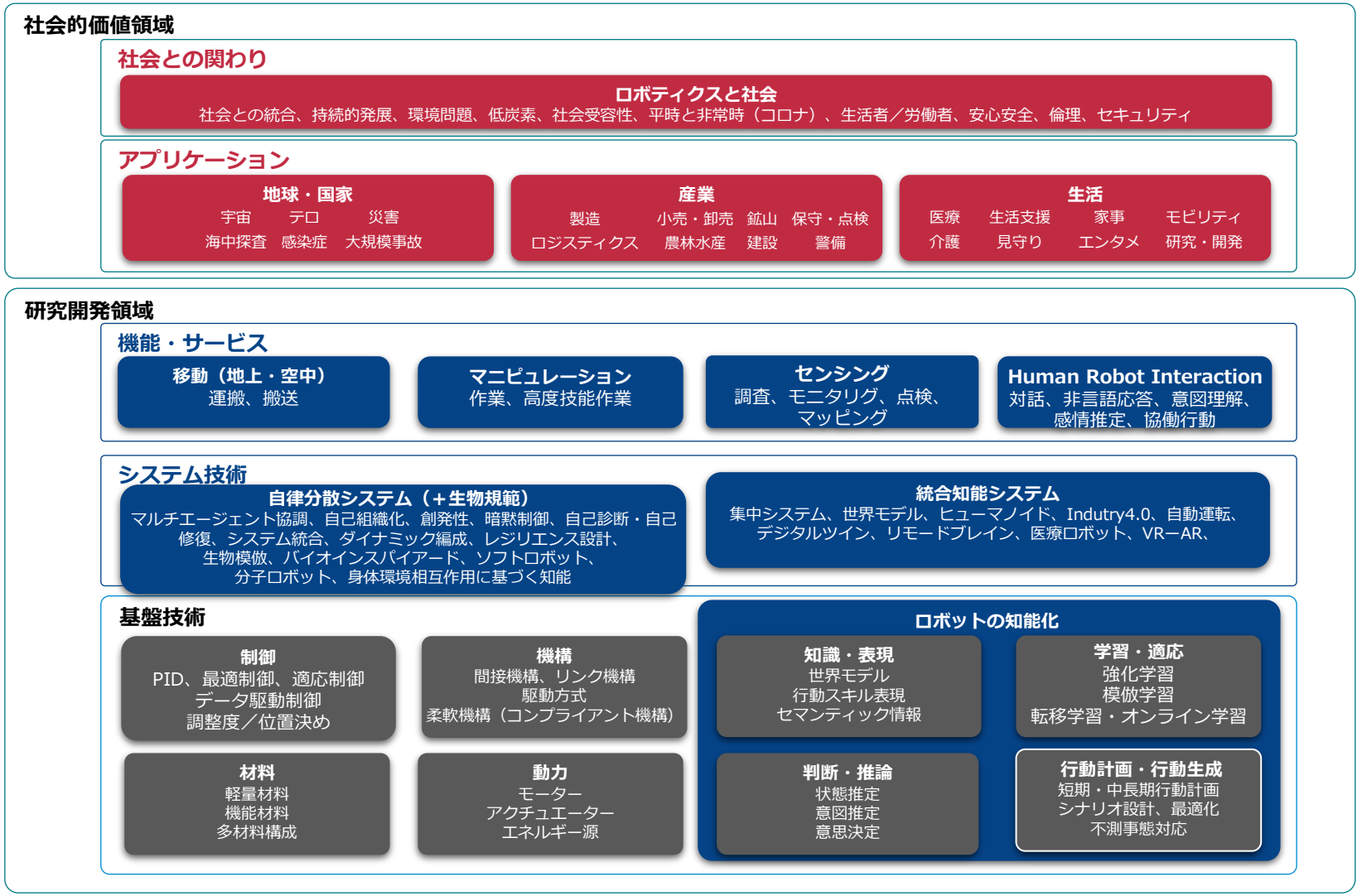
2. 遠隔・非接触技術としての再評価
COVID-19以降、医療支援、巡回、リモート接客など非接触・遠隔対応型ロボットの社会実装が加速。テレプレゼンス技術の応用拡大により、新たな人機協調の様式が形成されている。

3. 社会課題へのフィジカルAIロボットの適用
医療、災害、インフラ点検など多領域でフィジカルAIロボットの導入が進行。ロボット基盤モデルを活用した行動生成技術や、コンソーシアム型開発体制が国際的に拡大している。

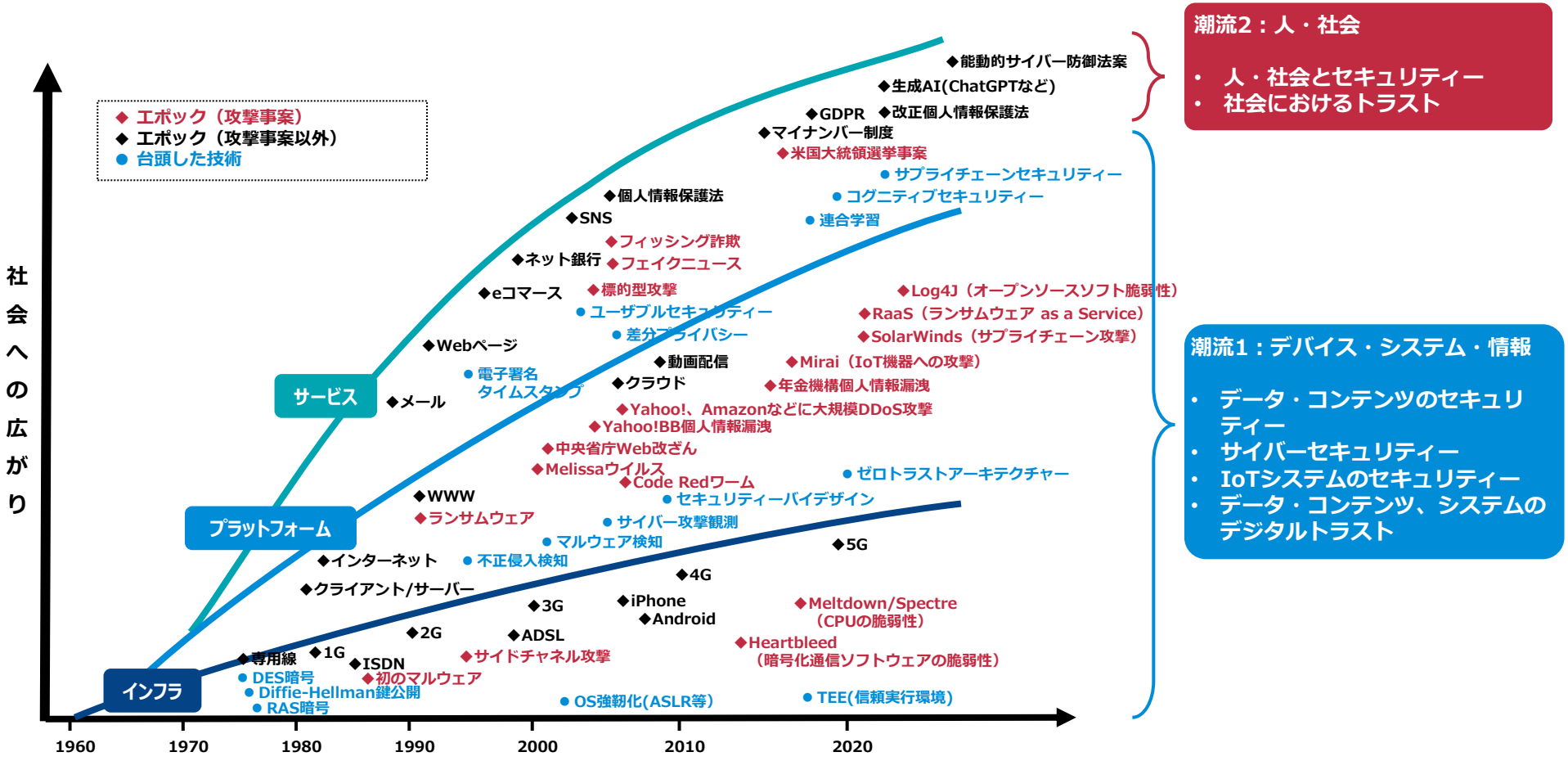
4. マルチモーダルAIによる知能アーキテクチャの革新
視覚・言語・行動を統合する大規模マルチモーダルモデルが本格化。ロボット特化型LLMや汎用基盤モデルにより、複雑環境への適応能力を備えた知能アーキテクチャが進化。

5. リアルタイム処理を支える次世代AIチップ
省エネルギー・高速処理・小型化を可能にする神経形態プロセッサやロボティクス向けエッジAIチップが実用段階へ。知能の分散化と即応性の両立を支える計算基盤として注目されている。

【ロボティクス区分】構造俯瞰図



【セキュリティー・トラスト区分】時系列俯瞰図



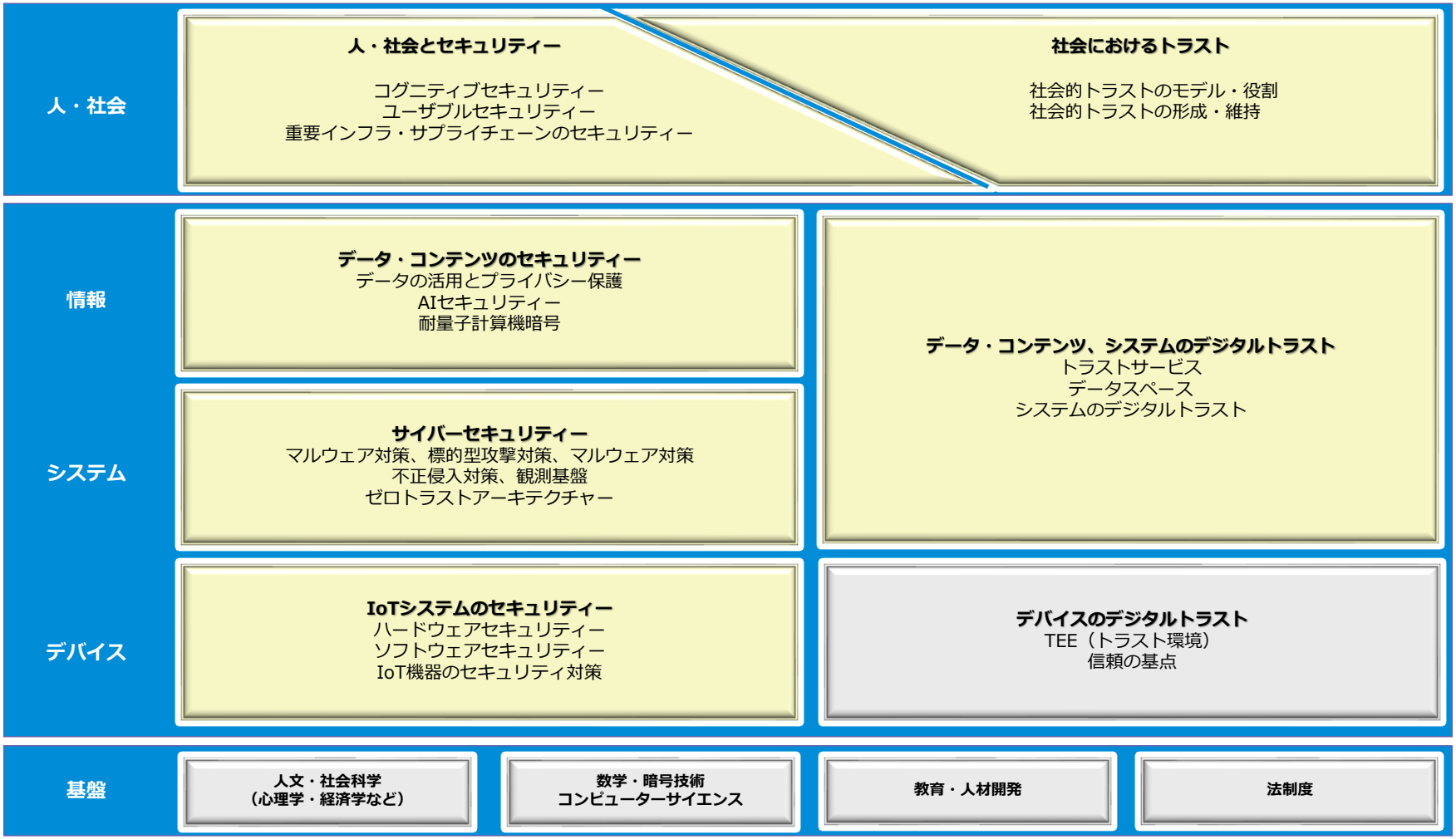
潮流2：人・社会

- 人・社会とセキュリティー
- 社会におけるトラスト

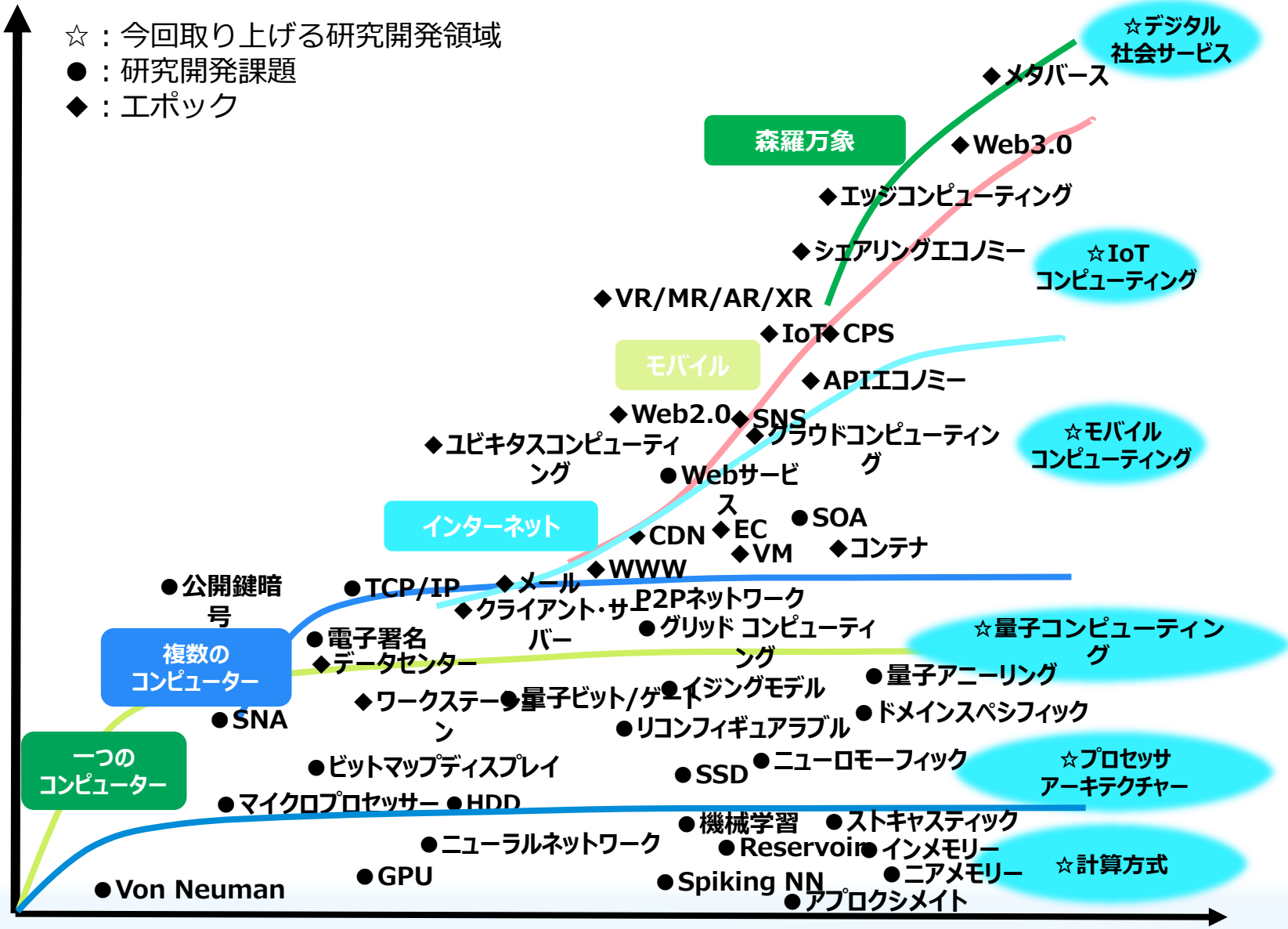
潮流1：デバイス・システム・情報

- データ・コンテンツのセキュリティー
- サイバーセキュリティー
- IoTシステムのセキュリティー
- データ・コンテンツ、システムのデジタルトラスト

【セキュリティー・トラスト区分】構造俯瞰図

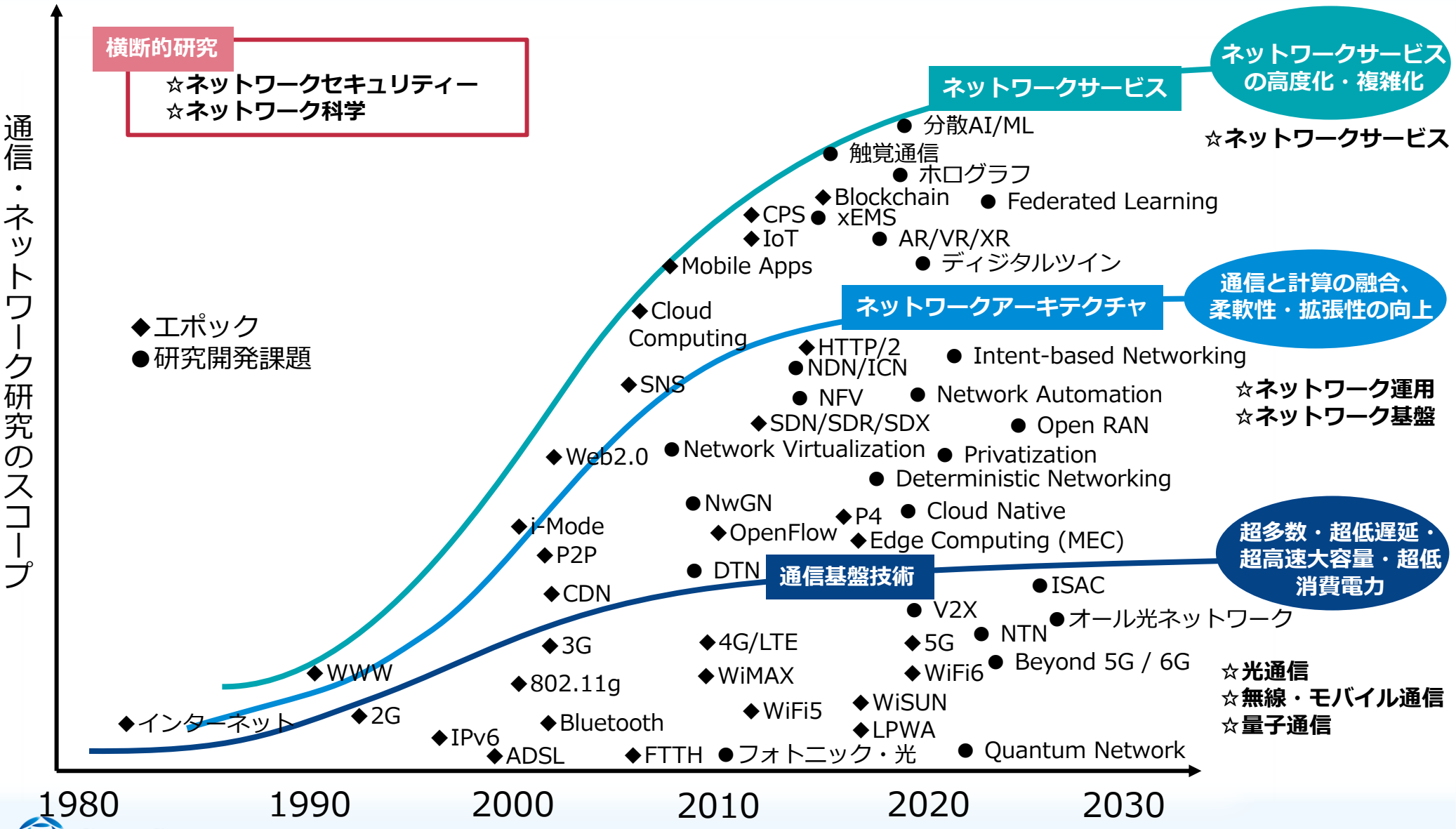


【コンピューティングアーキテクチャー区分】時系列俯瞰図



【コンピューティングアーキテクチャー区分】構造俯瞰図

【通信・ネットワーク区分】時系列俯瞰図



【通信・ネットワーク区分】構造俯瞰図

ネットワークサービス

⑥ネットワークサービス

サービスイネーブラー、オーケストレーター、情報指向ネットワーク（ICN）、サービス品質保証、高信頼ネットワーク、自己修復／予測保全、V2X、デジタルツイン

ネットワークアーキテクチャー

⑤ネットワーク運用

マルチレイヤーオーケストレーション、自律型ネットワーク、インテントベースネットワーク（IBN）、ネットワークテレメトリー、クラウドネイティブ

④ネットワーク基盤

Beyond 5G / 6G、オープン無線アクセスネットワーク（Open RAN）、RANインテリジェントコントローラー（RIC）、オール光ネットワーク（APN）、自動化・最適化、オープン化／ソフトウェア化／仮想化／プライベート化

通信基盤

①光通信

光ファイバー、光通信ネットワーク、フレキシブルグリッド、波長多重、空間多重、マルチバンド伝送、非地上系ネットワーク（NTN）

②無線・モバイル通信

Beyond 5G / 6G、ミリ波／テラヘルツ波、電波伝搬制御、アレーアンテナ、センシング通信融合（ISAC）、地上系・非地上系統合ネットワーク（TN-NTN）、時空間同期

③量子通信

量子鍵配送（QKD）、量子暗号ネットワーク、衛星量子暗号通信、量子セキュアクラウド、量子インターネット、量子中継、量子メモリー、量子テレポーテーション

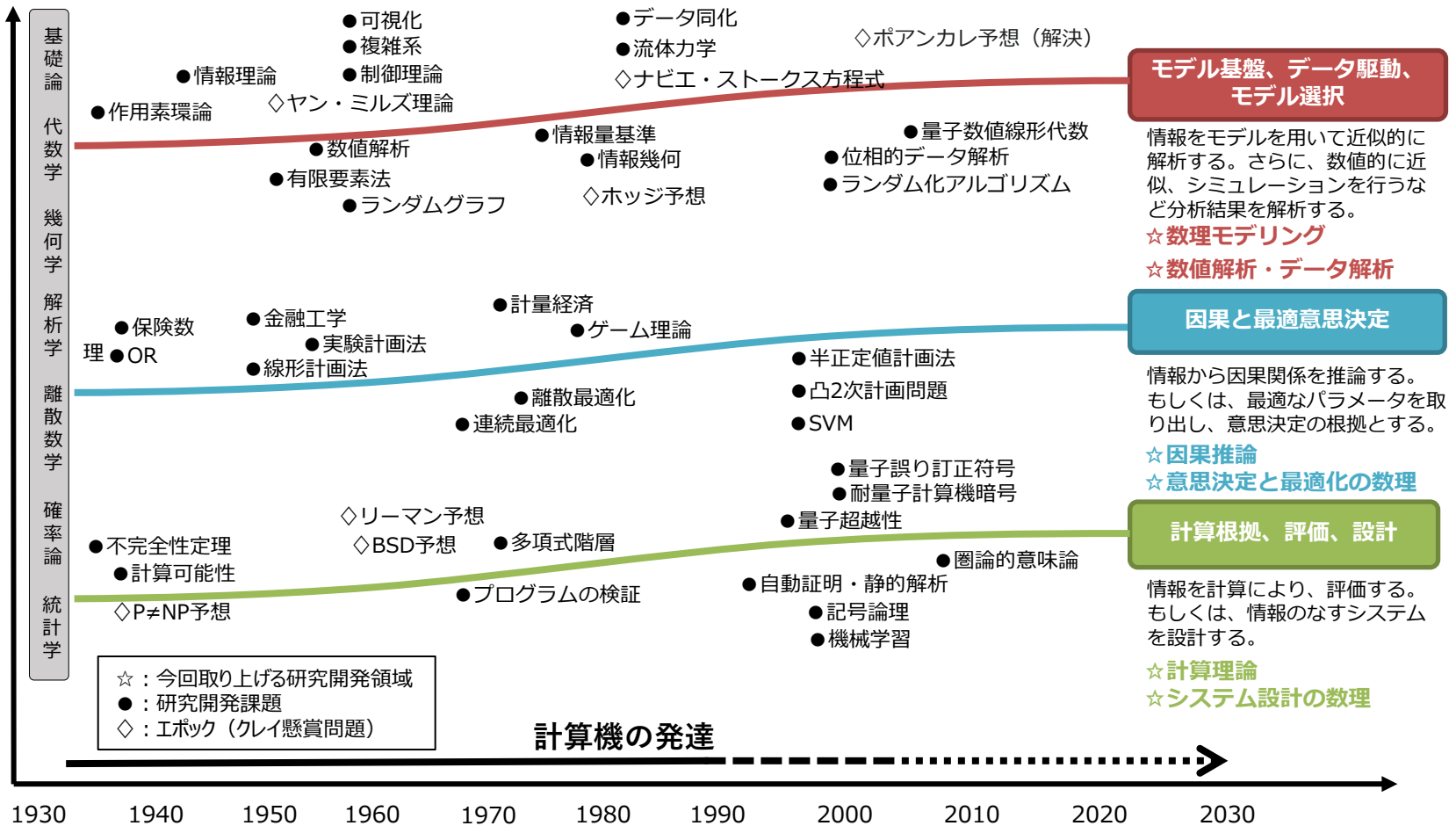
横断的研究

⑦ネットワーク科学

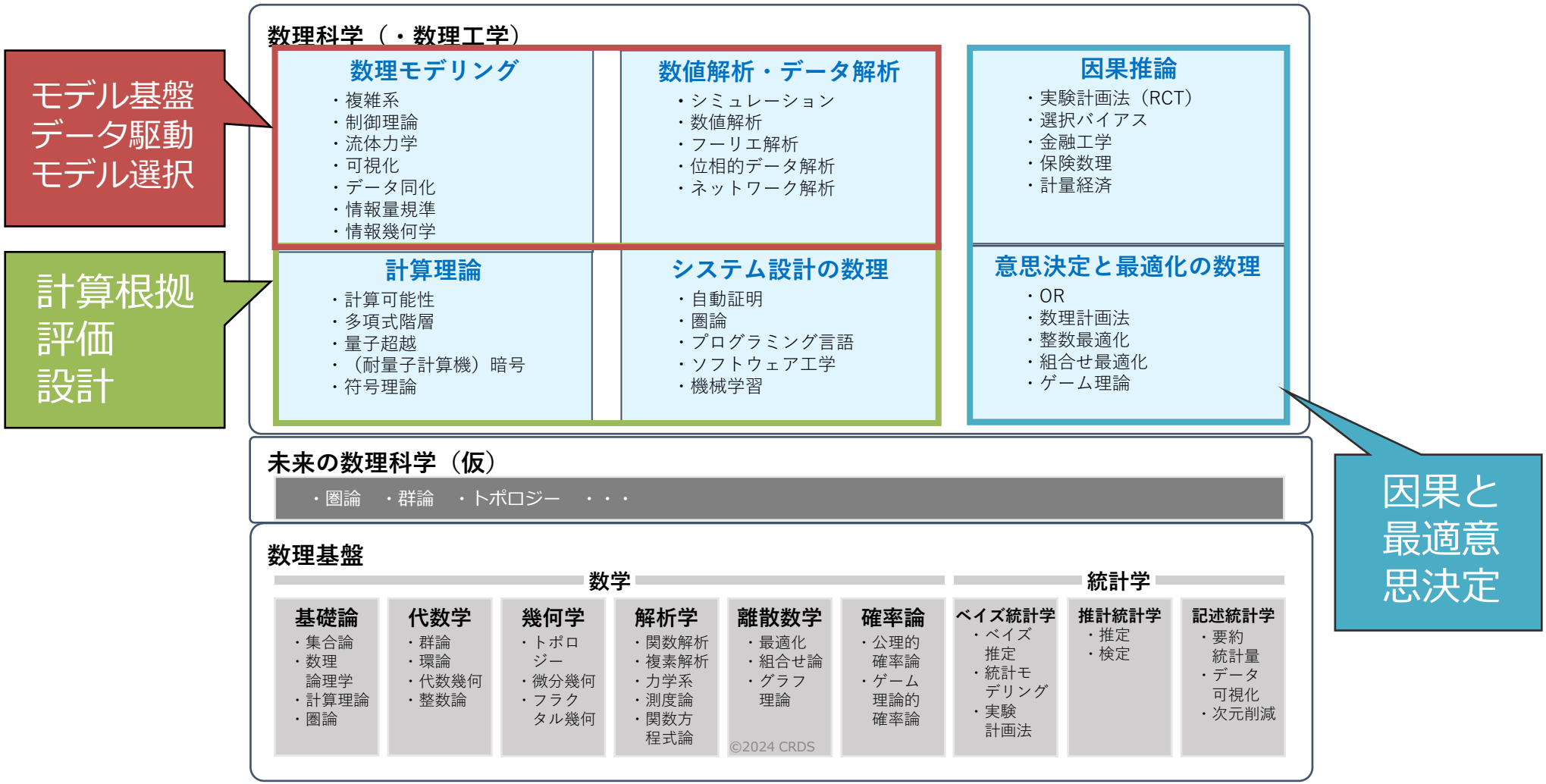
情報ネットワーク、ソーシャルネットワーク、複雑ネットワーク、高次相互作用ネットワーク、マルチレイヤーネットワーク、グラフニューラルネットワーク（GNN）、分子通信ネットワーク

ネットワーク セキュリティー

【数理科学区分】時系列俯瞰図



【数理科学区分】構造俯瞰図

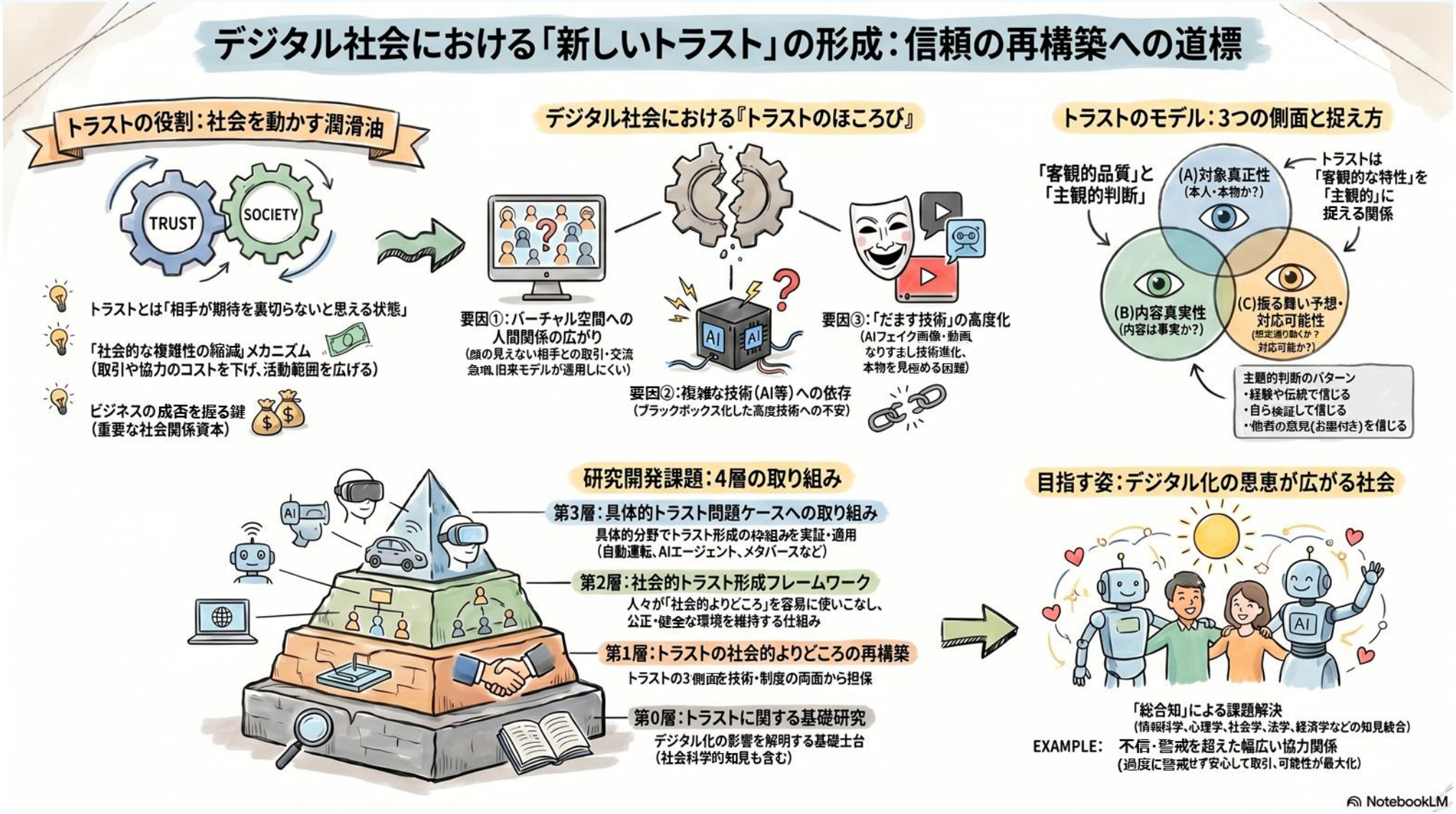


参考資料②
関連する戦略プロポーザル概要

俯瞰2026「わが国として重要な研究開発」と戦略プロポーザルの関係

		戦略プロポーザル					
俯瞰報告書2026版「わが国として重要な研究開発」		R3	R4	R5	R6	R7	R8(予定)
AIトランスフォーメーションと多層化するリスクへの対応	AI社会のリスク対策・トラスト基盤		デジタル社会のトラスト形成			戦略アップデート	AI社会のトラスト基盤
	コグニティブセキュリティー					コグニティブセキュリティー	統合
	AIトランスフォーメーション基盤	AI駆動科学			動向アップデート	[分野横断報告書]AI for Scienceの動向2026	
AIの自律化・フィジカル化・汎用化	次世代AIモデル	(R2: 第4世代AI)	戦略アップデート	次世代AIモデル	部分発展	フィジカルAIシステム	
	フィジカルAIシステム		リアルワールドロボティクス	戦略アップデート			
	バイオハイブリッドロボット						
	因果推論						
	最適化		情報・物理・数理の共創			意思決定を支援する最適化手法	
知能革命を支えるAIインフラ	通信と計算の融合				通信と計算の融合		AIコンピューティング
	AIコンピューティング						

R4年度 戦略プロポーザル の概要



AI社会のトラスト基盤：問題認識と研究開発課題

急速な技術発展、深刻化するAIリスクの最新状況を反映して、戦略提言(前頁)を更新・強化

AI社会のトラスト基盤：増大するリスクと研究開発の最前線

急速な先端AIの発展がもたらす深刻なリスクと、社会の信頼を守るための研究開発課題

深刻化するAIリスクの現状

1 **未知の脆弱性(ゼロディ)の大量発見**

先端AIモデル「Claude Mythos」がOSやブラウザの脆弱性を数千件発見。攻撃に悪用されれば社会インフラに甚大なダメージを与える恐れがあります。

2 **AGI/ASIによる制御不能のリスク**

人間の能力を超える汎用人工知能（AGI）や超知能（ASI）の出現が近づく中、人間による監視や制御がきかなくなる「暴走」のリスクが現実味を帯びています。

3 **認知戦(Cognitive Warfare)の巧妙化**

AIが人間になりすまして親密な関係を築き情報を盗み取ったり、思考を誘導したりする人間の認知面への攻撃が安全保障上の脅威となっています。

対策を阻む「3つの限界」

技術革新スピードへの遅れ

開発手法が演繹型（プログラミング）」から自律型（AIエージェント）」へ劇的にシフトしており、旧来の安全確保の方法論が通用しなくなっています。

断片的なアプローチ

本来トラストは「対象の真正性」「内容の真実性」「振る舞い予想可能性」の多層的なものですが、現状は特定の一面的な対策にとどまってます。

AI単体システム動作保証限界

AIは原理的に確率的な挙動をするため、100%の安全を事前保証することは不可能です。AIの外側に「ガードレール」や「トラスト基点」を組み合わせる枠組みが必要です。

取り組むべき3つの研究開発課題

①先端AIセーフティ

急速に拡大する新たな攻撃手法やリスクに対し、先端AIの振る舞いそのものを確実に制御し、完全性を確保するための革新的な研究を推進します。

②認知セキュリティ

技術的な客観面だけでなく、人間の主観的な判断（心理面）まで踏み込み、情報の真偽判断を撹乱する認知攻撃から社会を守る技術を開発します。

③統合トラスト基盤

AI単体に頼らず、社会全体でトラストを担保するためのフレームワークを構築、事後補償やリカバリー策、トラストアンカー（信頼の基点）の設置を含む総合的なフレームワークを目指します。

研究開発領域	
振る舞い予想・対応可能性	先端AIセーフティ
内容真実性	認知セキュリティ
対象真正性	統合トラスト基盤

NotebookLM

AI社会のトラスト基盤：社会的意義・戦略的意義

AI社会のトラスト基盤：安全・安心な共生社会への羅針盤

(a) 安全・安心な「人・AI共生社会」の実現

AIを単なる道具ではなく「自律的な主体」として捉え直し、暴走や悪用のリスクを極小化します。高度化するAIが社会インフラや市民生活に深刻なダメージを与える事態を未然防ぎ、信頼を基盤とした共生社会を構築します。

(b) 意思決定を支える「確かな情報源」の確保

AIを用いた高度な欺瞞(だまし)や思考誘導から、国民の自律的な意思決定を保護します。情報インフラ側での真正性保証(トラスト基点の確立)を強化することで、認知戦への耐性を高め、国家の安全保障に貢献します。

(c) 日本の強みとなる「総合的トラスト戦略」の確立

これまで各分野で個別に行われてきた「DFFT」「偽情報対策」「AI安全性」の取り組みを一つの枠組みへと統合します。「信頼の設計(Trust by Design)」そのものを日本の新たな国際競争力の源泉へと昇華させます。

トラストの総体(概念的枠組み)

[対象真正性＋内容真実性＋振る舞い予想・対応可能性] × [客観面＋主観面]

真の信頼構築には、多面的な要素の掛け合わせが必要不可欠です。

推進すべき3つの研究開発課題

トラストを社会全体で再構築するため、これら3つの研究開発を強力に推進します。

バラバラな施策を「統合」する日本の戦略

これらの施策群を「総合的トラスト戦略」のもとに有機的に集約・連携させることで、日本発の世界をリードする信頼のグランドデザインを描き、国際社会におけるルール形成を主導します。

NotebookLM

概要：コグニティブセキュリティー

コグニティブセキュリティーとは、**「デジタル社会において、脆弱性となりうる人間の認知に悪影響を及ぼす情報から人・社会を守ること」** である。

生成AI、SNSなどの情報技術の進展により、フィッシングやフェイクニュースなど、人間の認知に悪影響を及ぼす情報による問題は、個人、組織にとどまらず社会、国家にまで拡大している。

これらの問題に対処するために、**情報を受け取る人間の認知を守り、自律的な意思決定を支援するコグニティブセキュリティーが重要**となっている。

コグニティブセキュリティーにおける自律的な意思決定とは、悪意のある情報や偏った情報に騙されたり依存したりして、本人が望まない結果をもたらすことのない意思決定であり、人間の認知に悪影響を及ぼす情報による問題に対して、**自律的な意思決定を行うことを支えるためにコグニティブセキュリティーの研究開発が必要不可欠**である。

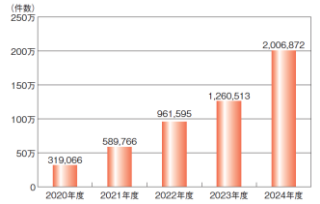
問題の例

フィッシング

- フィッシング報告件数は、約171万件（令和6年）となり120万件（令和5年）と比較すると約1.4倍に増加。
- フィッシングメールの約8割がAIを利用しているという調査報告もあり、手口が巧妙化している。

偽・誤情報

- 能登半島地震の際に救助要請を求める偽情報が拡散され救助活動に支障をきたした。
- 米国防総省近くで爆発”偽画像で株価が一時下落した。
- 投降呼びかけるゼレンスキー氏の偽動画による情報工作が行われた。



年度別フィッシング報告件数

（出典）IPA情報セキュリティ白書2025 図1-1-6

“米国防総省近くで爆発”偽画像で株価一時下落する騒動に

出典）総務省「デジタル時代における放送制度の在り方に関する検討会（第19回）NHK説明資料」
https://www.soumu.go.jp/main_content/000884978.pdf

能登半島地震での偽の救助要請による活動妨害

出典）NHK NEWS WEB「「不謹慎で迷惑」能登半島地震で相次いだ偽救助要請 実態は?」
https://www3.nhk.or.jp/news/html/20240312/k10014383261000.html

投降呼びかけるゼレンスキー氏の偽動画

出典）Youtube
https://www.youtube.com/watch?v=X17yrEV5sl4

研究開発の内容

本戦略プロポーザルは、情報技術が進展したデジタル社会において自律的な意思決定を支援するコグニティブセキュリティーの研究開発課題と推進方法を提言する。

自律的な意思決定を支援するコグニティブセキュリティーのために、事前対策を中心とした現在の現在対策（①）から、情報を受け取る人間の認知特性の知見に基づく新たな対策手法を創出して、即時対策と事後対策を含む広範な対策（②）の実現を目指す。

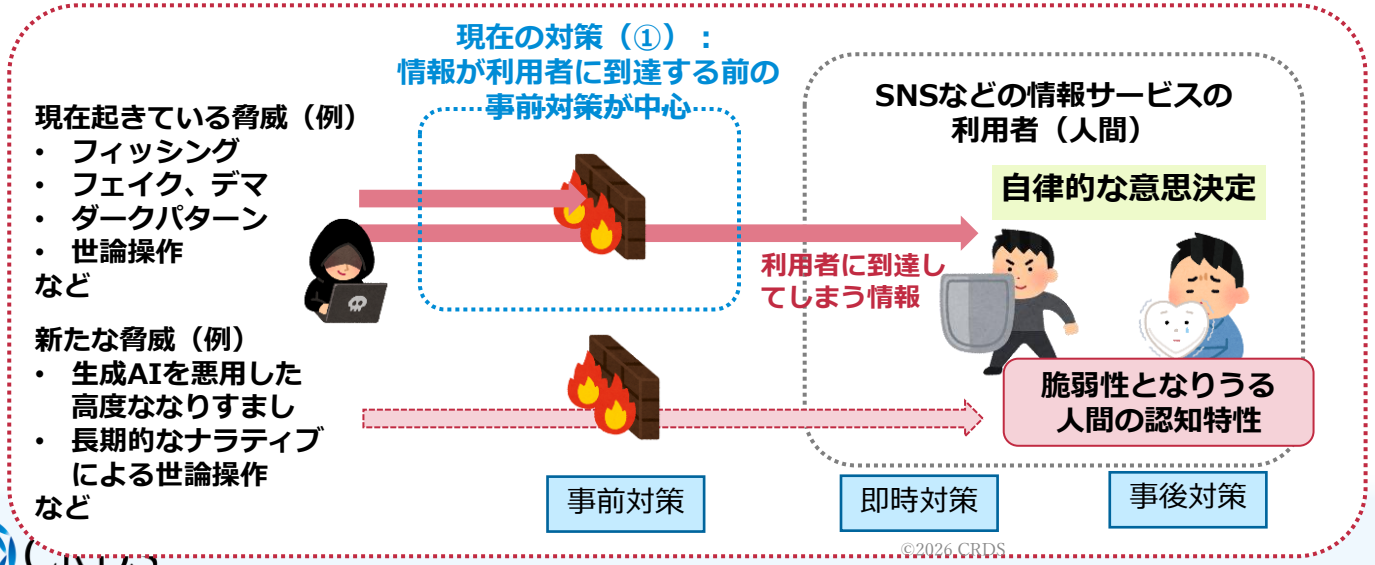
現状認識

- 情報技術の進展によりフィッシング詐欺やSNS上の偽・誤情報による詐欺、デマ、世論操作など、情報による問題（脅威）が拡大
- 現状、脅威につながる情報が利用者（人間）に到達する前の対策（事前対策）が中心

解決すべき問題点

- **利用者（人間）に到達してしまう情報への即時対策、事後対策を含む広範な対策が十分にできていない**
- **対策が後追いになっており、将来の新たな脅威への対策が十分にできていない**

本提言で目指す対策（②）：事前対策に加えて即時対策と事後対策を含む広範な対策



具体的な研究開発課題

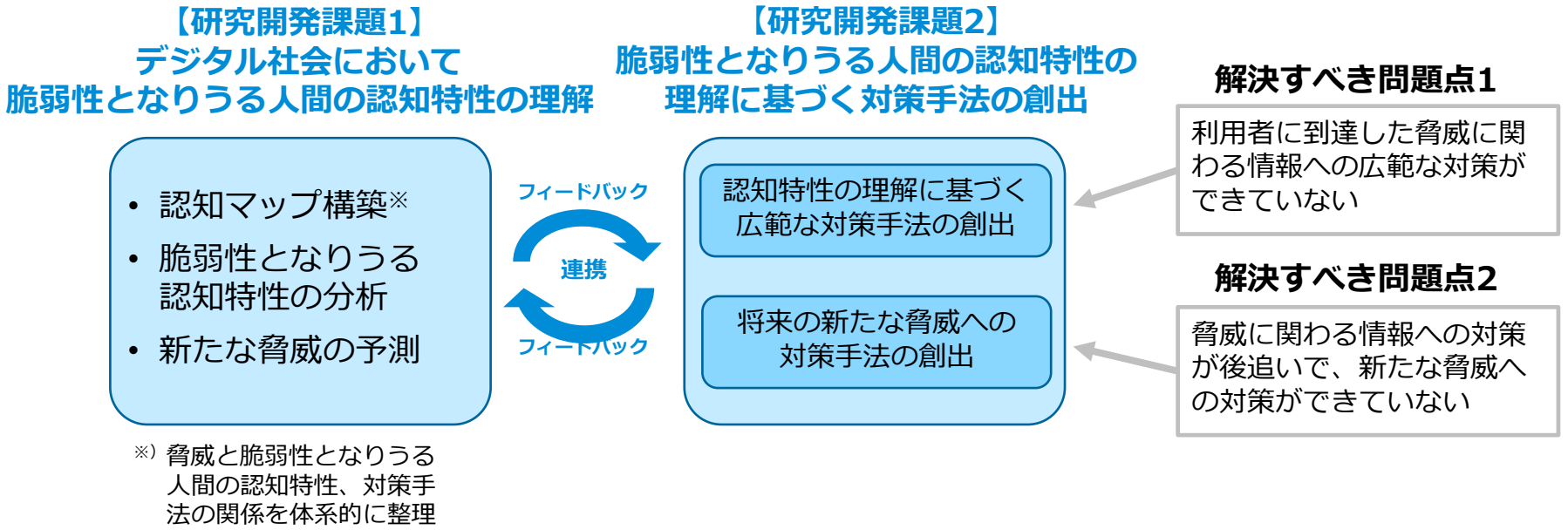
研究開発課題1：
デジタル社会において脆弱性となりうる人間の認知特性の理解

連携

研究開発課題2：
脆弱性となりうる人間の認知特性に基づく対策手法の創出

具体的な研究開発課題

「解決すべき問題点1、2」を解決するために、デジタル社会において脆弱性となりうる人間の認知特性の理解と、その理解に基づく対策手法を創出する2つの研究開発課題を提言する。



※）脅威と脆弱性となりうる人間の認知特性、対策手法の関係を体系的に整理

解決すべき問題点1に対する取り組み

研究開発課題1において、脅威と脆弱性となりうる人間の認知特性、対策手法の関係を体系的に整理した認知マップを構築し、狙われている認知特性を特定・分析する。研究開発課題2において認知特性の分析結果を活用して対策手法を創出する。

解決すべき問題点2に対する取り組み

研究開発課題1において、認知マップからまだ狙われていない認知特性を特定・分析し、新たな脅威を予測する。
研究開発課題2において、新たな脅威への対策手法を創出する。

概要：フィジカルAIシステム

戦略プロポーザル概要：フィジカルAIシステムの研究開発 ～身体性を備えたAIとロボティクスの融合～

フィジカルAIシステムとは？

身体性を備えたAIへの進化

サイバー空間のAI（頭脳）

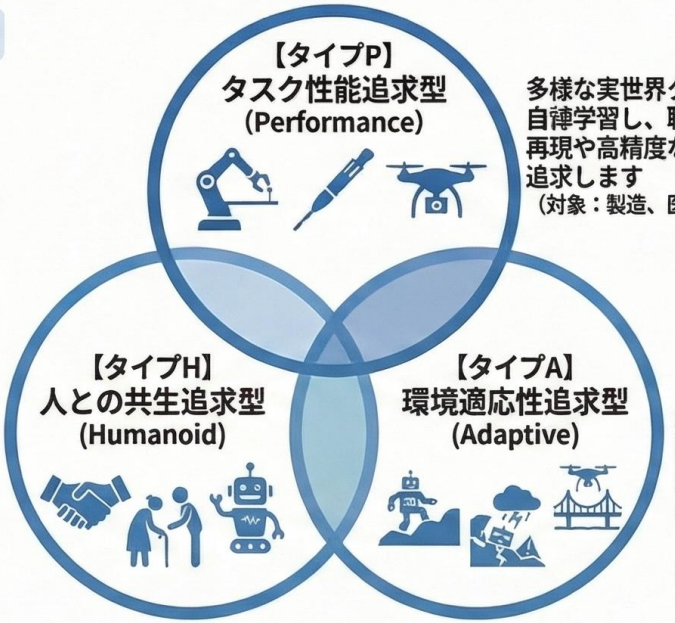
物理的な身体（ロボット）

実世界で自律的に行動・学習

実世界での相互作用が鍵

センサーからの入力と動作結果に基づき、環境に適応しながら知能を獲得・発達

AIロボットが目指す「3つの価値特性」



【タイプP】タスク性能追求型（Performance）
多様な実世界タスクを自律学習し、職人技の再現や高精度な作業を追求します
（対象：製造、医療、点検）

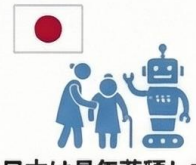
【タイプH】人との共生追求型（Humanoid）
人の意図や文脈を理解し、協働・共進化する自然な共生環境を構築します
（対象：介護、生活支援、教育）

【タイプA】環境適応性追求型（Adaptive）
遍在・変動・未知の環境でも安定して移動するロバスト性を備えます
（対象：農業、災害対応、インフラ点検）

日本の勝ち筋（JAPAN STRATEGY）

米中の量産戦術 vs 日本の現場適応

米中の物量作戦



日本は最年蓄積した「制御・総合技術」と「現場の知見」で対抗

「分野特化型」のアプローチ

汎用性のみを追わず、具体的な社会課題に即した「現場ですり合わせる力」を最大化

4つの研究開発課題

1. 実世界適応型フィジカルAIの開発

エッジでのリアルタイム実行や低消費電力を実現し、ハードとAIを高度に統合

2. 身体性を備えた知能の解明

身体構造と知能の関係を多角的に研究し、認知発達の視点から知能獲得プロセスを解明

3. 人に安全なシステムの構築

設計段階から安全性を取り入れる「Safety-by-Design」により、不確実な状況でも信頼性を確保

4. 社会的影響（ELSI）と制度設計

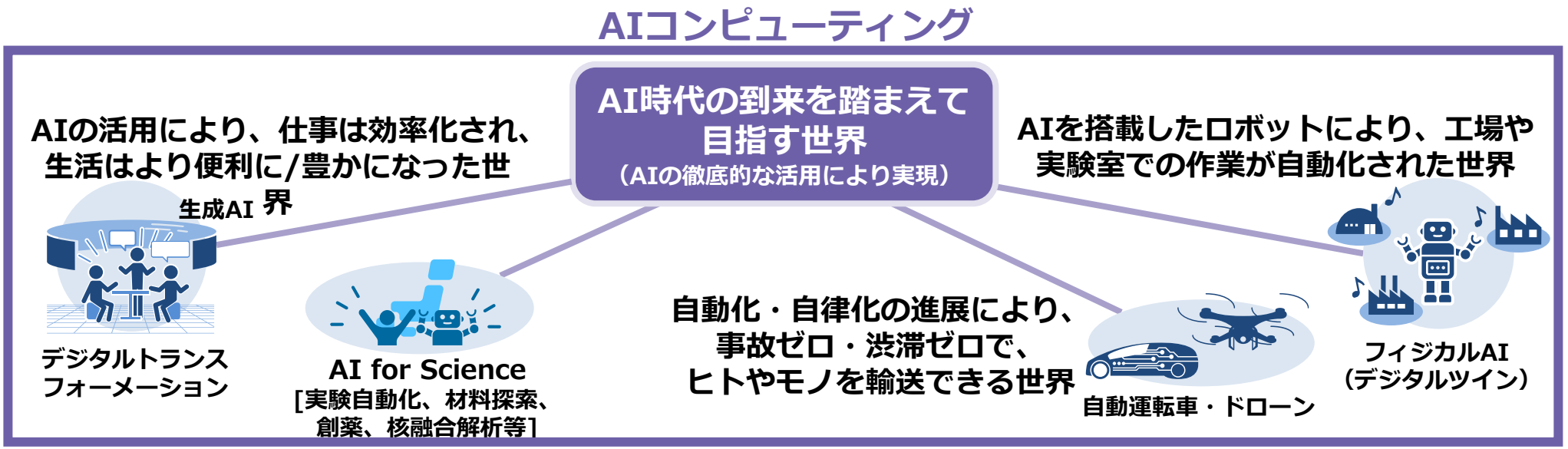
法的責任や倫理的課題を整理し、技術開発と並行して社会に受容されるための制度を整える

NotebookLM

概要：AIコンピューティング

AI時代の到来により、社会インフラとしての**「計算基盤」**が必要不可欠になってきた

- 「ビッグデータ時代以上に、**大量のデータを使うことで価値を生み出す**時代」
- 「誰もが、いつでも、どこでも、**AIを活用する**時代」・・・**AI処理の民主化**



目標

大量の計算を“短時間”に“低消費電力”で実行する コンピューティング技術の実現

目標を実現するにあたり、様々な観点を踏まえつつ高速化と省電力化を進める必要がある

■技術的観点

- 専用性/汎用性とスケーラビリティー（特定用途向けの小型化や、大規模AIモデルの汎用化など）
- 開発環境や運用管理基盤との整合性（設計ツール、開発ライブラリ、運用管理ソフトなど）
- 数理的・物理的な裏付け（AI向けの確率・近似理論、半導体物理、材料科学など）

■経済的観点

- 量産性と耐故障性を踏まえたコスト効率の向上
- 経済安全保障を踏まえたサプライチェーンの確立

■社会的観点

- カーボンニュートラル（システムの導入から、運用、さらに廃棄に至るまでのライフサイクル）

問題点と研究開発課題（案）

問題点

1. AIの進展により、処理すべきデータが膨大となり、計算が追い付かない
   ⇒ 高速化が必要

2. 高速化しようとすると、継続的な運用が困難なほど消費電力が増大する
   ⇒ 省電力化が必要

研究開発課題（案）

①「計算資源の連続体（データセンターとエッジをまたがる計算基盤）」を想定した、分散コンピューティング＆ネットワークアーキテクチャーの研究開発

②データセンタースケールのコンピューティングシステムの研究開発

③高速性と省電力性をバランスさせたプロセッサー※の研究開発
※演算用半導体チップ

④数理や物理を基礎とする計算・通信方式の研究開発

研究開発課題を設定するにあたり前提とするAIコンピューティングの将来像予想

- データセンターは必要不可欠な社会インフラのひとつとなる
- データセンターとエッジとが高度に融合し（計算資源の連続体[Computing Continuum]となり）、常に何をどこで処理するかを動的に選択することになる
- 計算機の小型・高性能化が進み、実世界（物理世界）に散らばるハードウェアを有効活用する機運が高まる