

「米国人工知能研究開発戦略計画」について

2016年10月25日
NEDO ワシントン事務所

2016年10月12日、大統領府は「米国人工知能研究開発戦略計画（National Artificial Intelligence Research and Development Strategic Plan）」¹を発表した。

1. 本計画の目的及びアウトカム

AIは、多大な社会的・経済的利益を生み出す可能性のある革新的技術であり、生活、仕事、学習、コミュニケーションのあり方が変わる可能性がある。AI研究は、経済発展の推進、教育機会とQOLの増進、国家安全保障の強化など、優先事項の推進にも役立つと考えられている。このため、米国政府はこれまでにAI研究に投資し、2015年の米国政府による投資額は約11億ドルである。

連邦政府による投資が画期的な発見につながってきた一方、民間でもAIに多大の投資を行っている。このように投資状況が変化していることから、AI技術開発に向けた連邦政府投資の適切な役割、官民・国際協力の機会を検討する必要がある。

このため、大統領府国家科学技術会議（NSTC）のネットワークング・IT研究開発小委員会（NITRD）は、AIタスクフォースを立ち上げ、特に産業界単独では担うことが困難な分野を中心に、AIのR&Dにおける優先事項を策定した。本計画では、AIについて戦略研究目標を定義し、民間が投資する可能性が低い分野に連邦予算を集中させ、併せてR&D人材の拡大・維持に対応するための優先事項を取りまとめている。

本計画では、AIにおける科学技術ギャップの特定と、そのギャップを埋めるための連邦政府のR&Dにおけるフレームワークを定義しているほか、AIへの短期～長期的支援における優先事項を特定している。各省庁の研究内容を指定するのではなく、行政府全体の目標を設定し、その中で各省庁が、ミッションや能力、予算等に適合した形で優先事項に対応し、全体としての研究ポートフォリオが戦略計画に則ったものとなることを意図している。

2. AIの現状

2013年から2015年の間、「deep learning」に言及した科学雑誌記事数は6倍となっている。また、研究のグローバル化が進んでおり、論文数で見ると、米国を抜いて中国が世界1位となっている。

AI研究においては政府が重要な役割を果たしてきているが、民間においても主に短期的収益が見込まれる分野での関連R&Dが活発である。「deep learning」や「deep neural net」という言葉を含む特許数が急増しているほか、AIベンチャーに対するベンチャーキャピタルの投資は、2013年から2014年で4倍に膨れ上がっている。

¹ https://www.whitehouse.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf

3. R&D 戦略

以下は、本戦略計画の全体構成を図化したものである。

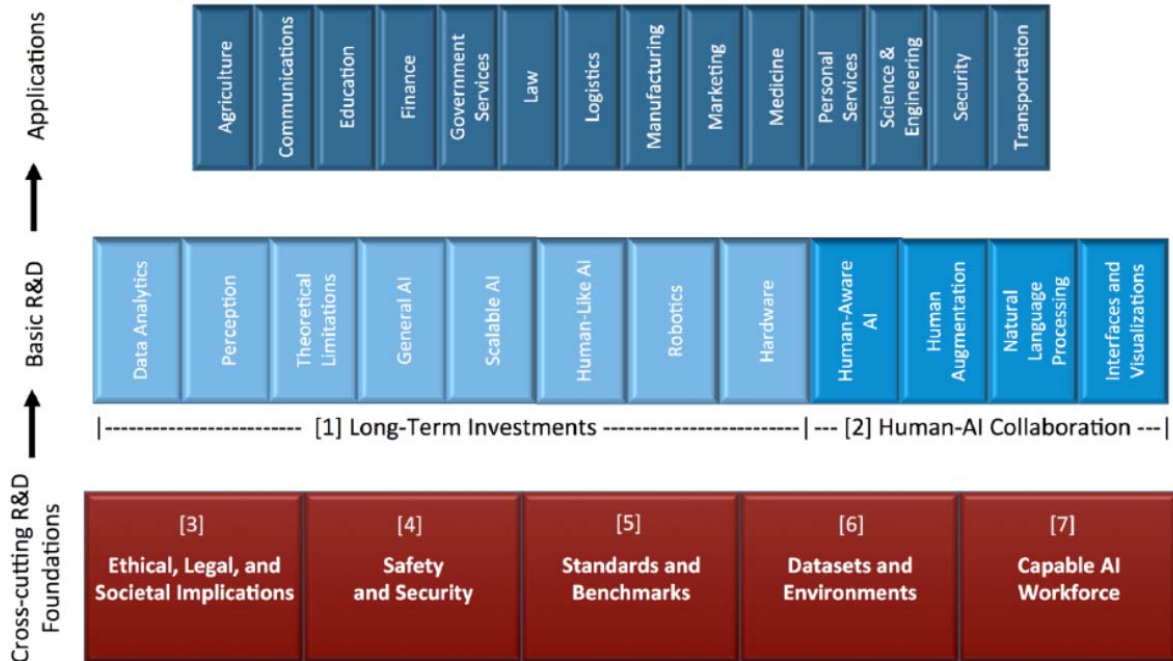


Figure 4. Organization of the AI R&D Strategic Plan. A combination of crosscutting R&D foundations (in the lower red row) are important for all AI research. Many basic AI R&D areas (in lighter and medium dark blue row) can build upon these crosscutting foundations to impact a wide array of societal applications (in top dark blue row). (The small numbers in brackets indicate the number of the Strategy in this plan that further develops each topic. The ordering of these Strategies does not indicate a priority of importance.)

前述した目標を達成するために、本計画では、連邦政府の優先事項として以下を特定した。

<戦略 1：AI 研究への長期投資>

5～10 年以上先に高価値の結果を生むためには、ハイリスク研究への長期的な投資が必要である。重要な研究分野は、以下のとおり。

- **データに着目した知識発見手法の促進**：インテリジェントデータの理解や知識発見を達成するための基礎的なツール・技術。ビッグデータに隠れている有益な情報を特定できるマシン・ラーニングアルゴリズムの開発、データクリーニング手法の効率性向上、データの不統一性や異常を発見する手法の創出、人からのフィードバックを統合するアプローチの開発、データと関連メタデータの同時マイニング手法の探索等。
- **AI システムの知覚能力の増強**：より強靱で信頼できる知覚を可能とするハードウェア及びアルゴリズム、より遠距離にあるデータをリアルタイム・高解像度で獲得できるセンサー、多彩なセンサーと他のソース（クラウド等）を統合して、AI システムが知覚し、将来を予測できるようにするための知覚システム、知覚プロセスにおける不確実性を計算・反映するフレームワーク等。
- **AI の理論上の能力及び制限の理解**：AI の理論上の能力及び制限について理解が不足しているため、AI システムの動きをよりよく理解するための理論研究が必要。数学、コンピュータ科学等の

様々な学問分野が関係するが、AI システムパフォーマンスを理解するための統一的な理論モデルやフレームワークがない状況。

- **汎用 AI に関する研究推進**：特定タスクを行う「Narrow（特化）AI」（IBM Watson や DeepMind 社の AlphaGo 等）ではなく、「General（汎用）AI」に向けた研究を進める。実現には今後数十年はかかるとみられており、長期的・継続的な研究が必要。
- **拡張性のある AI システムの開発**：複数の AI システムや人間の連携による計画、制御、協力を可能とするための、効率的・強靱・拡張可能な手法の発見。
- **人間的な AI の研究の増進**：効率的に人間を補佐する次世代システムの実現には自らを人間が理解できる方法で説明できる人間的な AI が必要。現行の AI と人間の間での学習・行動には大きな違いがあるため、新たなアプローチ開発に向けた基礎的研究が必要。
- **より高性能で信頼性の高いロボットの開発**：リアルタイムで様々なセンサーから得られる情報を抽出するロボット知覚の理解の増進、物理的世界をよりよく理解して行動できるような認知・推論の向上、ロボットの動きを向上させるためのモビリティ・操作分野の研究、信頼性が高く予測可能な形でシームレスに人間と協力できるロボットの開発。
- **高性能 AI のためのハードウェア改善**：AI アルゴリズムに最適化されたハードウェアの開発。
- **ハードウェア改善のための AI 開発**：AI による高性能コンピュータの性能予測や最適化が行われつつあるが、AI システムを用いた新たなアプローチによりハードウェアの性能向上が可能。

<戦略 2：人・AI 協力に向けた効果的な手法の開発>

人の代替ではなく、人と協力して最適なパフォーマンスを達成する AI システムを開発する。人と AI システムの間の効果的な相互作用を創出するための研究が必要であり、システムデザインが過剰な複雑性、信頼不足や過信につながらないための R&D を行うべき。以下のような課題が存在。

- **人間との協力を前提とした AI のための新しいアルゴリズムの開発**：ユーザーと直感的にやり取りでき、シームレスな協力を可能とする AI。
- **より良い視覚化手法及び AI ヒューマンインターフェースの開発**：人間が大量のデータや情報をよりよく理解。
- **より効果的な言語処理システムの開発**：実世界に近い環境での言語処理を可能とする研究や、人と AI システムの間のインタラクションをより直感的・自然にすることを旨とする。

<戦略 3：AI の倫理・法的・社会的意義の理解・対処>

AI 技術は、人に適用される公式・非公式な規範に従って行動することが期待されている。AI の倫理的・法的・社会的な位置付けを理解し、それらの原則に沿った AI システムをデザインするための手法を開発するための研究が必要。コンピュータ科学、社会・行動学科学、倫理、バイオ医療科学、心理学、経済学、法学、政策学といった学際的取り組みが関連。以下は研究課題の代表例である。

- **AI デザインによる公平性・透明性・アカウントビリティの本質的な向上**：AI における適切なデータの収集及び利用のため、デザイン段階で組み込む。
- **倫理的な AI の構築**：適切な価値観を反映したトレーニング用データセットを創出するには、領域横断的アプローチが必要。また、実世界の複雑な状況に対応できるように、価値観ベースの適切な矛盾を解決。

- **倫理的思考を統合した AI のアーキテクチャのデザイン**：監視エージェントの組込等、AI システムの全体的デザインを倫理・法律・社会的目標に適合させるための研究。

<戦略 4：AI システムの安全性・セキュリティの確保>

AI システムが大規模普及する前に、安全、セキュアかつ管理された手法でシステムが運用されることが保証されなければならない。研究課題は以下のとおり。

- **説明性・透明性の向上**：AI がなぜその結論に至ったかの理由をユーザーに対して説明できるような基本能力と透明性を持つシステムの開発。
- **信頼の構築**：有益な情報を提供でき使いやすいインタフェースを備えた、正確で信頼できるシステムの開発。
- **AI システムの検証・妥当性確認のための新手法**：ロバストで信頼性を持つためには、自己評価、自己診断、自己修復ができる能力の開発。
- **攻撃に対する防御**：事故に対応できるだけのロバスト性を持ちつつ、様々な種類のサイバー攻撃に対してもセキュアである必要。
- **長期的安全性・価値観整合の実現**：AI システムは自己修正を重ねていくことから、安全性を保つためにも、元々の目標に一致した活動が行われているかをチェックするための自己モニタリングアーキテクチャ、評価中のシステムの開放を防ぐ制限戦略、価値ラーニング、および、自己修正に影響されない価値フレームワーク等の開発。

<戦略 5：AI トレーニング・試験のための共有公共データセット・環境の開発>

トレーニング用データセットやリソースの深さ、質、正確性が、AI の性能に大きく影響する。高品質なデータセットと環境を開発し、またデータに含まれる商業的・個人的権利を尊重しつつアクセス者を限定しないための研究が必要である。以下が重要な研究分野である。

- **AI の関心及びアプリケーションの多様性に対応できるような多様なデータセットの開発・共有**：政府が既に有する既存のデータセットや連邦政府予算で整備できるデータセット、及び、可能な範囲で、産業が持つデータセットを広く利用可能とする技術。
- **民間や公共の関心に対応したトレーニング・試験リソースの開発**：重要な研究課題としては、データキャプチャ、キュレーション、分析、視覚化など。価値ある知識を多量のデータから抽出するために必要な科学が遅れている。
- **オープンソースソフトウェアライブラリやツールキットの開発**：連邦政府は、オープン AI 技術の開発・支援・利用に対する取り組みを増進することで、イノベーションの継続的推進を支援できるほか、政府内でオープン AI 技術の利用を増やすことで、開発者に対する参入障壁を低減。

<戦略 6：標準・ベンチマークを通じた AI 技術の測定・評価>

標準、ベンチマーク、テストベッドと、それらが実際に AI コミュニティに利用される事が AI 研究の促進・方向付けに不可欠。以下が推進すべき分野。

- **広範囲にわたる AI 標準の開発**：AI の実用化・普及拡大に遅れないよう、標準開発も加速する必要。ソフトウェアエンジニアリング、パフォーマンス、メトリクス、安全性、ユーザビリティ、相互運用性、セキュリティ、プライバシー、トレーサビリティ、ドメインが、標準において対応すべき分野。

- **AI 技術のベンチマークの確立**：AI 技術を効果的に評価するため、適切かつ効果的な試験手法や定量化手法の開発、標準化。
- **利用可能な AI テストベッドの増加**：特定の連邦政府内に設けられたセキュアなテストベッド環境を整備し、大学・民間研究者に開放することも想定。
- **標準・ベンチマークへの AI コミュニティの参画**：標準化を進め、その産学官における普及を推進するためには政府のリーダーシップ及び調整。

＜戦略 7：米国における R&D 人材ニーズの理解増進＞

AI の R&D 推進においては、AI 研究者の強いコミュニティが必要となる。AI 分野の現在及び将来の R&D 人材需要に対する理解を向上させることが、本計画に記載された戦略 R&D 分野に対応するための AI 専門家の確保に必要。これには教育・再トレーニング機会、多様性などの知見も含まれるべき。

4. 提言

提言 1：科学技術上の機会を特定し、AI への R&D 投資の効果的な調整を支援するための R&D 実施フレームワークを、上記の 1～6 の戦略に対応する形で構築。

連邦省庁は、本計画の中で示された R&D 課題に対する R&D 実施フレームワークを立ち上げるべきである。これにより、各省庁の計画・調整・協力が容易になる。実施フレームワークは、各省庁の優先事項、能力、権限、予算を考慮に入れたものでなければならない。

本戦略計画の実施を支援するため、NITRD は、AI に焦点を当てた省庁横断型ワーキンググループを立ち上げることを検討すべきである。

提言 2：本計画の戦略 7 に沿う形で、大量の R&D 人材を創出・維持するために国内状況を調査。

本報告書に提示された R&D 戦略課題に対応するために、大量かつ活力のある AI R&D 人材が重要となる。AI の研究者・開発者が今後不足するとの報告もあるが、人材の現状や需給状況に関する公的なデータは存在しない。NITRD は、現行及び将来の R&D 人材ニーズの特定方法を調査し、ニーズに対応できる R&D 人材を確保するための調査または提言を策定すべきである。