

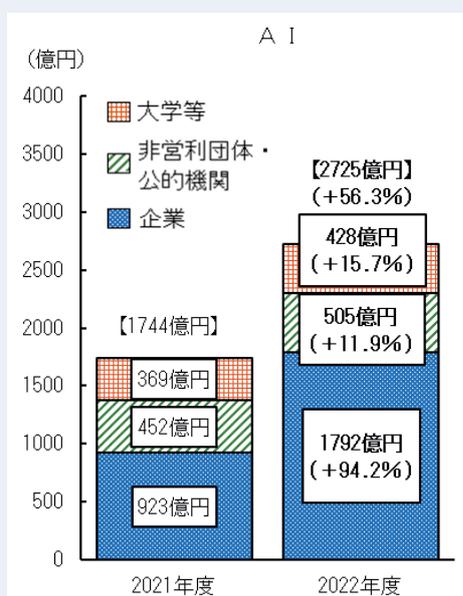
第2章 我が国におけるAI関連研究開発の取組

前章で見てきたような近年の急速なAI技術の進展や国際的な議論を踏まえ、新たに設置された「AI戦略会議」において、「AIに関する暫定的な論点整理¹」が令和5年(2023年)5月に取りまとめられ、政府では、AIに関する国際的な議論と多様なリスクへの対応、AIの最適な利用、AI開発力の強化等に向けた取組を進めてきています。また、研究機関や大学、

企業等において、AIに関する様々な研究開発や社会実装に向けた取組が進められてきており、当該分野における研究開発費も増加してきています(第1-2-1図)。

この章では、我が国におけるAI研究開発におけるこれまでの主な経緯を振り返りつつ、第3次ブーム以降の主な取組を紹介します。

■ 第1-2-1図 / AI分野の研究主体別研究費



資料：総務省統計局「令和5年科学技術研究調査結果²」

第1節 我が国におけるAI関連研究開発の歩みと近年の取組

我が国の研究者も、第1次ブームの段階から、AI技術の進展に貢献してきています。1960年前後から、我が国の大学でも自然言語処理、音声認識、画像処理等の研究開発が始まり、1970

年代には日本語情報処理の本格的な研究が始まるなど、AI研究が広がりました。

そして、昭和57年(1982年)から大型プロジェクトとして始まった通商産業省(当時)の

1 AI戦略会議(令和5年)「AIに関する暫定的な論点整理」 https://www8.cao.go.jp/cstp/ai/ai_senryaku/2kai/ronten.pdf



2 総務省統計局「令和5年科学技術研究調査結果」 https://www.stat.go.jp/data/kagaku/kekka/kekkgai/pdf/2023ke_gai.pdf

「第5世代コンピュータプロジェクト」等を通じて、第2次ブームを迎え、大学や民間企業で多くの人材が育てられましたが、実用化の限界が意識され、ブームも終息しました。

その後、2010年代以降の深層学習をはじめとする技術の進展を受けて、再び、AIの活用可能性への関心が高まり、平成29年(2017年)3月に「人工知能技術戦略」、平成31年(2019年)3月に「人間中心のAI社会原則」、令和4年(2022年)4月に「AI戦略2022」がそれぞれ取りまとめられました。これらの戦略等に基づき、総務省、文部科学省、経済産業省をはじめとする関係府省や国立研究開発法人等が連携しながら、AI技術の研究開発や社会実装に向けた取組が進められるとともに、民間企業やスタートアップ企業においても研究開発が進められてきました。主な取組を以下に紹介します。

●産業技術総合研究所人工知能研究センター¹の設立

平成27年(2015年)5月に設置された同センターでは、これまでAIの要素機能の研究開発で多数の成果を挙げ、使いやすい形のプログラムに実装したソフトウェアモジュールを構築・公開し、「生産性の向上」、「健康、医療・介護」、「空間の移動」などの広範な分野で応用技術を開拓してきています。実世界にAIを埋め込んでいくために更に必要な基盤技術に焦点を当て、人間と協調できるAI、実世界で信頼できるAI、容易に構築できるAIの三つの

柱の下、基礎研究を社会実装につなげるための研究開発を進めています。

●理化学研究所革新知能統合研究センター(AIPセンター)²の設立

平成28年(2016年)4月に設置された同センターでは、世界最先端の研究者を糾合し、革新的な基盤技術の研究開発や我が国の強みであるビッグデータを活用した研究開発を推進しています。汎用基盤の研究については、深層学習の原理の解明、現在のAI技術では対応できない高度で複雑・不完全なデータ等に適用可能な基盤技術の実現を目指しています。また、我が国の強みをAIの活用で伸ばしつつ、AIを用いた社会課題の解決を指向しています。加えて、AIと人間の関係としての倫理の明確化や、AIを生かす法制度の検討などを行っています。

●情報通信研究機構知能科学融合研究開発推進センター³の設立

平成29年(2017年)4月、従来から情報通信研究機構が蓄積してきたデータを含め、産学官が利用しやすい形での研究開発環境を整備するとともに、知能科学領域における次世代研究開発を推進するオープンイノベーション型の戦略的な研究開発推進拠点として設置されました。

1 産業技術総合研究所人工知能研究センター
<https://www.airc.aist.go.jp/>



2 理化学研究所革新知能統合研究センター
<https://www.riken.jp/research/labs/aip/>



3 情報通信研究機構知能科学融合研究開発推進センター
<https://www2.nict.go.jp/ais/>



●株式会社Preferred Networksによる深層学習フレームワーク「Chainer」の開発

同社は、AI研究の最前線で活動している、平成26年(2014年)に設立されたスタートアップ企業です。同社は独自の深層学習フレームワーク「Chainer」を開発して、平成27年(2015年)6月にオープンソースソフトウェアとして公開しました¹。複雑なニューラルネットワークを直感的かつ柔軟に構築できることから、研究者や開発者から多くの支持を受け、深層学習に関する技術開発の黎明期の発展に寄与しました。なお、令和元年(2019年)12月から、同社は米国Meta社と連携して「PyTorch」の開発に参加し、PyTorchエコシステムに貢献しています²。

●トヨタ自動車株式会社によるAI研究開発

トヨタ自動車株式会社は、平成28年(2016年)1月に、米国シリコンバレーに、AIの研究開発を行う新会社トヨタ・リサーチ・インスティテュート(TRI)を設立しました。TRIの使命は、新しいツールや能力を開発し、

人々の能力を拡張して生活の質を向上させることです。モビリティの革新を推進するため、TRIはロボット工学、人間中心のAI、相互作用的な運転、エネルギーと素材の分野で世界トップクラスのチームを構築しています。

●人工知能研究開発ネットワーク(AI Japan R&D Network)³の設立

人工知能研究開発ネットワークは、我が国の英知を糾合し、AI研究開発の活性化を図ることを目的に、令和元年(2019年)12月、産業技術総合研究所、理化学研究所、情報通信研究機構が中核となり、AIの研究開発などに積極的に取り組む大学や公的研究機関などから構成されるコンソーシアムとして設立され、AIの研究開発に関する連携が進められてきました。その後、令和5年(2023年)4月に、任意団体として新たに設立され、さらに民間企業も会員に加わり、AIに関する研究開発や成果利用の促進等に係る更なる連携が進められています。

1 株式会社Preferred Networks「Deep Learning Frameworks」 <https://www.preferred.jp/ja/projects/dlf/>

2 株式会社Preferred Networks「Preferred Networks、深層学習の研究開発基盤をPyTorchに移行」
<https://www.preferred.jp/ja/news/pr20191205/>

3 人工知能研究開発ネットワーク <https://www.ai-japan.go.jp/>



コラム1-2 AI研究第一線の方に聞くAIブームとこれから

●岡野原大輔氏

株式会社Preferred Networks 代表取締役 最高研究責任者



AI研究の第一線を走っていらっしゃる株式会社Preferred Networksの岡野原代表取締役にお話を伺いました。

AIとは何か、の定義については、著名な研究者の定義でも60~70の定義が存在し、世の中に明確な定義はない、そしてそれは世界的に共通の認識です。

現在のAIブームは、2022年11月のChatGPTリリースが牽引していることは間違いありません。それまでは特定のタスク向けのAIが多かったものの、こうした基盤モデルは様々な目的に使用可能で、必ずしもAIの専門家ではない一般の人が容易に使えるものであるため、ユーザー数が急激に伸びました。今後もAI市場は拡大していきます。

一方で、過去にはAIに対する過剰な期待と幻滅が定期的に続いてきたので、今回もそうしたことが起こる可能性はあります。ハルシネーション(幻覚)は大きな問題になり得ます。現時点では、影響がクリティカルな領域では最終的に人が見張っておくことが必要と言えるでしょう。

今後のAI発展の戦略について、分野特有の知識を必要とする分野へ、汎用基盤モデルを特化、細分化していくことで、我が国が競争力を持つことができると考えています。特に、材料探索、生命科学、気象、宇宙といった分野の知識が必要な領域は相性が良いです。ほかにも数学やプログラムも相性が良く、これらの分野ではブレークスルーが起きる直前だと思っています。

AIの科学や研究開発への影響について、コンピュータが登場した時に研究の姿が変貌したのと同様に、研究の姿は変わるでしょう。例えば、研究において人間が見つけられない対称性やパターンを見つけれられるようになるかもしれません。

AIの研究について、うまくいかない場合も多くあり、後ろ指をさされることもあると思います。過去には実際に難しかったものを今研究する場合にも批判があり得るでしょう。しかし、AIは指数関数的に性能が上がっており、今やってみたらうまくいくことがあるかもしれません。日本の中でもある意味で「空気を読まずに」研究できる環境であると良いと望んでいます。

第2節 我が国における生成AIに関する研究開発について

OpenAI社のChatGPTをはじめ世界的に大規模言語モデルの開発が進む中、日本語を扱う能力の高いモデルが少ない¹といった課題や、一部の企業による独占への懸念が存在します。こうした状況を受け、我が国でも、大学や国立研究開発法人、企業等において、日常生活・産業現場での活用も想定した、高度な日本語処理が可能な日本語大規模言語モデルや軽量版モデル等の生成AIの開発が急速に進められています。

日本語に特化した大規模言語モデルの開発

における先駆的な取組として挙げられるのは、令和2年(2020年)の東京大学発スタートアップ、株式会社ELYZAによる「ELYZA Brain」の開発です。その後、同社は令和3年(2021年)8月、入力したテキストデータを3行に要約する「生成型」の要約モデル「ELYZA Digest」を一般公開し²、さらに令和4年(2022年)3月には、キーワードから約6秒で日本語の文章を生成できる大規模言語モデル「ELYZA Pencil」を一般公開しました³。また、令和6年(2024

1 GPT-3における日本語データの割合は0.11%

参考: https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv

2 株式会社ELYZA 「どんな文章でも3行にできる要約AI “ELYZA DIGEST”、公開5日間で13万人が利用」
<https://prtimes.jp/main/html/rd/p/000000012.000047565.html>

3 株式会社ELYZA 「国内初。キーワードから約6秒で文章生成ができる 日本語の文章執筆AI “ELYZA Pencil” を一般公開」
<https://prtimes.jp/main/html/rd/p/000000015.000047565.html>

※なお、上述の「ELYZA Digest トライアル版」及び「ELYZA Pencil トライアル版」のサービスは終了しており、「ELYZA LLM for JP (デモ版)」にて類似機能が利用できる。

年) 3月には、パラメータ数が700億パラメータ(70B)の日本語大規模言語モデル「ELYZA-japanese-llama-2-70b」を開発し、グローバルモデルに匹敵する性能を実現したことを発表しています¹。

また、令和5年(2023年)12月、東京工業大学情報理工学院情報工学系の岡崎直観教授と横田理央教授らの研究チームと産業技術総合研究所は、日本語能力に優れた生成AIの基盤である大規模言語モデル「Swallow」を公開しました²。これは、英語の言語理解や対話で高い能力を持つオープンソースの大規模言語モデル(米Meta社Llama2)の日本語能力を拡張することにより開発されました。現在、パラメータ数が70億パラメータ(7B)、130億パラメータ(13B)、700億パラメータ(70B)であるモデルが公開されており、オープンで商用利用も可能となっています³。

さらに、令和5年(2023年)5月に始動した、情報・システム研究機構国立情報学研究所が主宰する「LLM勉強会(LLM-jp)⁴」では、オープンかつ日本語に強い世界トップレベルのLLMの構築の開発を目指しており、その第一歩として、10月には130億パラメータ(13B)であるLLMの構築・公開を果たしています。また、現在、1,750億のパラメータ数を持つ大規模言語モデルの構築に取り組んでいます。

民間企業においても様々な取組が進められており、例えば、日本電気株式会社(NEC)は、130億パラメータの高い日本語性能を有する軽量大規模言語モデル「cotomi」を開発し、令和5年(2023年)8月から商用を開始する

とともに⁵、小型～大規模モデル(パラメータ数1,000億クラス)の開発も進められています。

ソフトバンク株式会社は、生成AI開発向けの計算基盤を整備するとともに(第3節参照)、その子会社であるSB Intuitions株式会社は、この計算基盤を活用して、日本語に特化した国産LLMの開発を開始しています。今後、3,900億パラメータの国産LLMの令和6年度(2024年度)内の構築を目指しています⁶。

日本電信電話株式会社(NTT)は、学習時に用いる変数の規模を表すパラメータ数が70億パラメータの「軽量版」と、同6億パラメータの「超軽量版」の二つのモデルからなる、軽量で日本語性能の高い大規模言語モデル「tsuzumi」を開発し⁷、令和6年(2024年)3月に商用化を発表しました⁸。tsuzumiは文字のほか文書に含む図や表、グラフの内容を読み取って回答するマルチモーダルな機能も備えており、今後は視覚情報に加えて、音声のニュアンス・顔の表情・ユーザーの置かれている状況などのモーダル拡張に対応する予定となっています。また、tsuzumiの応用例の一つ

■ 第1-2-2図 / tsuzumiとロボットの連携の様子



提供：日本電信電話株式会社

1 株式会社ELYZA「ELYZA、グローバルモデルに匹敵する日本語LLMを開発、デモ公開」

<https://prtimes.jp/main/html/rd/p/000000042.000047565.html>

2 TokyoTech-LLM「Swallow」<https://tokyotech-llm.github.io/swallow-llama>

3 産業技術総合研究所「日本語に強い大規模言語モデル『Swallow』を公開」

https://www.aist.go.jp/aist_j/press_release/pr2023/pr20231219/pr20231219.html

4 LLM勉強会 <https://llm-jp.nii.ac.jp/>

5 日本電気株式会社「NEC、日本市場向け生成AIを開発・提供開始～業種ナレッジの構築を目指したカスタマープログラムを開始～」

https://jpn.nec.com/press/202307/20230706_01.html

6 ソフトバンク株式会社「2024年3月期 第3四半期 決算説明会」

https://www.softbank.jp/corp/set/data/ir/documents/presentations/fy2023/results/pdf/sbkk_earnings_presentation_20240207.pdf

7 日本電信電話株式会社「NTT 独自の大規模言語モデル『tsuzumi』を用いた商用サービスを2024年3月提供開始」

<https://group.ntt.jp/newsrelease/2023/11/01/pdf/231101aa.pdf>

8 NTTコミュニケーションズ株式会社「NTT 版LLM『tsuzumi』を活用したソリューションの提供～同時にパートナーシッププログラムの募集を開始～」https://www.ntt.com/content/dam/nttcom/hq/jp/about-us/press-releases/pdf/2024/0325_2.pdf

として、身体感覚を持つロボットと連携し、物理的作業の制御についても取り組んでいます（第1-2-2図）。

また、Google社出身の研究者等が令和5年（2023年）7月に東京に設立したAIスタートアップSakana AI株式会社は、進化や集合知等の自然界の原理を応用した基盤モデルの開発を目指して、令和6年（2024年）3月、既存のオープンソースモデルを融合（マージ）して、より高度な新しい基盤モデルを自動的に作成する技術「進化的モデルマージ」を開発したことを発表しました¹。

主要諸外国でも大規模言語モデルの開発が進められ、国際競争が加速する中、経済産業省では、国内の生成AIの開発力強化を目的とし、生成AIのコア技術である基盤モデルの開発

に対する計算資源の提供支援や、関係者間の連携促進、対外発信等を実施するプロジェクト（GENIAC²）を開始しました。その中で、新エネルギー・産業技術総合開発機構の「ポスト5G情報通信システム基盤強化研究開発事業」において、競争力があり波及効果大きい基盤モデルの開発を行う企業等を公募のうえ令和6年（2024年）2月に第1-2-3表の7法人等を選定し³、採択された法人等に対する計算資源の確保と利用料補助という形での支援を開始しました。また、開発者同士のネットワークや生成AIの利活用促進のため、セミナーや開発者ネットワークイベント、開発者・利用者のマッチングイベント等も実施しています。

■ 第1-2-3表 / 「ポスト5G情報通信システム基盤強化研究開発事業」に選定された法人等一覧

実施者	概要
株式会社ABEJA	LLMの社会実装に向けた特化型モデルの元となる汎化的LLMを研究開発
株式会社Preferred Elements	100B/1Tパラメータからなる大規模マルチモーダル基盤モデルを構築
東京大学	多様な日本語能力の向上を目指した公開型の基盤モデルを開発
Sakana AI株式会社	大きな計算リソースを必要とする基盤モデルを高効率化するための、重要な技術（蒸留、MoE、強化学習）を用いたモデルを開発
ストックマーク株式会社	実ビジネス活用において懸念・障害となるハルシネーション（もっともらしい嘘）を抑制した基盤モデルを開発
情報・システム研究機構国立情報学研究所	オープンかつ日本語に強いGPT-3級（1,750億パラメータ）の大規模言語モデルを構築
Turing株式会社	完全自動運転を見据え、日本の運転環境に強い適応力を持つ、マルチモーダル（言語・画像・映像）基盤モデルを開発

資料：経済産業省作成の「採択事業テーマ概要」を基に文部科学省作成

1 Sakana AI株式会社「進化的アルゴリズムによる基盤モデルの構築」
<https://sakana.ai/evolutionary-model-merge-jp/>

2 Generative AI Accelerator Challenge
https://www.meti.go.jp/policy/mono_info_service/geniac/index.html

3 経済産業省「『ポスト5G情報通信システム基盤強化研究開発事業』の採択事業者を決定しました（令和6年2月2日）」
https://www.meti.go.jp/policy/mono_info_service/joho/post5g/20240202.html

第3節 基盤モデルの開発を支える計算資源やデータ資源の整備や活用

大規模言語モデル等の基盤モデルの開発・構築には、大量の学習用データや計算資源が必要となり、そのためのコストや環境負荷も課題となっています。このため、研究機関と産業界との計算資源の共有、計算資源の開発や利用環境整備の取組に対する支援が進められています。

「富岳」に代表される我が国のスーパーコンピュータは、AI技術の開発や深層学習の計算にも活用されており、複雑なシミュレーションや解析を行う際の強力なツールとして利用されています。富岳において、AIやデータサイエンスを活用した取組を重点分野として研究開発が推進されるなど、AIとスーパーコンピュータの組合せにより、これまでにないスケールのデータ解析やモデル学習が可能になってきています。

令和5年(2023年)5月には、東京工業大学、東北大学、富士通株式会社、理化学研究所は、富岳を活用して、超大規模な並列計算環境において大規模言語モデル学習を効率良く実行する技術の開発に着手しました^{1,2}。今回の研究開発

の成果物は令和6年度(2024年度)の公開が予定されており、我が国の多くの研究者やエンジニアによる活用を通じて、次世代の革新的な研究やビジネスの成果につながることを期待されています。また、理化学研究所では、超伝導方式の国産量子コンピュータ初号機「叡」と富岳との連携に関する研究開発も進めています。

また、令和6年(2024年)4月から稼働した東京工業大学学術国際情報センターの最新スーパーコンピュータ「TSUBAME4.0」は、高性能のGPUを搭載したことなどにより、現存する国内のスーパーコンピュータの中では、富岳に次ぐ2位相当の演算性能を持ち、科学技術計算・ビッグデータ解析・AIなど幅広い分野で積極的に活用されることが期待されています(第1-2-4図)。

また、産業技術総合研究所が構築・運用するAI処理向け計算インフラストラクチャ「AI橋渡しクラウド(ABC I³)⁴」(第1-2-5図)では、生成AIの基盤的な開発力強化を目的とした設備の拡張を速やかに進めるとも

■ 第1-2-4図 / 東京工業大学学術国際情報センターの最新スーパーコンピュータ「TSUBAME4.0」



提供：東京工業大学

■ 第1-2-5図 / ABC Iの外観



提供：産業技術総合研究所

- 1 理化学研究所「スーパーコンピュータ『富岳』政策対応枠における大規模言語モデル分散並列学習手法の開発について」
https://www.riken.jp/pr/news/2023/20230522_2/index.html
- 2 令和5年8月より、株式会社サイバーエージェント、名古屋大学、Kotoba Technologies Inc.が参画機関に追加。
- 3 AI Bridging Cloud Infrastructure
- 4 ABC I <https://abciai/ja>



■ 第1-2-6表 / 大規模言語モデル構築支援プログラム2023採択課題

	機関名	課題名	時期	ノード数
第1回	株式会社Preferred Networks	日本語データを含めた大規模データによる大規模言語モデル学習	8月-9月期	60
第2回	代表機関：情報・システム研究機構国立情報学研究所 参加機関：産業技術総合研究所、東京工業大学、LLM-jp	オープンソースかつ日本語に強いGPT-3級（1,750億パラメータ）の大規模言語モデルの構築とそれに関連する研究開発の推進	10月-11月期	60
	株式会社ELYZA	オープンな英語圏LLMの日本語化と性能改善	12月-1月期	40

資料：産業技術総合研究所ウェブサイトを基に文部科学省作成

に、大規模言語モデルの構築を目指す利用者が、高性能かつ最大80ノード¹（640GPU）にも及び豊富な計算資源を最大60日にわたって予約して活用できる「大規模言語モデル構築支援プログラム²」を提供しました。第1-2-6表のとおり、令和5年度には3機関の取組が採択され、第2節で紹介した日本語に強い大規模言語モデル「Swallow」や「ELYZA-japanese-Llama-2-70b」の開発にもつながりました。

また、クラウドサービスがあらゆる国民生活・産業活動にとって不可欠となる中、基盤的なクラウドサービスの提供に用いられるプログラム（基盤クラウドプログラム）を開発する

ための産業基盤を早急に国内に確保する必要があることから、経済安全保障推進法に基づき、安定供給確保を図るべき特定重要物資として「クラウドプログラム」を指定し、重要な技術開発や高度な計算機の利用環境整備に取り組む事業者の計画を認定した上で、必要な経費の補助をしています。これまで、東京大学（量子コンピュータを活用したクラウドサービス提供）、さくらインターネット株式会社、ソフトバンク株式会社、株式会社ゼウレカ（生成AIを含む基盤クラウドプログラム向けの計算資源のクラウド提供）等の計画を認定しています（第1-2-7図、第1-2-8図）³。

■ 第1-2-7図 / さくらインターネットデータセンターの設備



提供：さくらインターネット株式会社

■ 第1-2-8図 / ソフトバンクの生成AI計算基盤の設備



提供：ソフトバンク株式会社

1 ノードとは「結び目」や「節」を意味する単語で、スーパーコンピュータ分野では一つの管理単位をノードと呼ぶことが多い。例えば、一つの基本ソフト（OS）が動作しているCPUやメモリの塊を指す。
 2 産業技術総合研究所「大規模言語モデル構築支援プログラム」https://abciai/ja/link/llm_support_program.html
 3 経済産業省「クラウドプログラムの安定供給の確保」https://www.meti.go.jp/policy/economy/economic_security/cloud/index.html

また、AIモデルの開発には多様な学習が必要になります。このため、内閣府は、政府等が保有するデータについて、AI開発者からのニーズに応じてAI学習データとしての提供を促進するため、令和5年（2023年）11月、「AI学習データの提供促進に向けたアクションプランver1.0」を策定しています¹。

OpenAI社のGPTモデルなどの大規模言語モデルは、ウェブから集められた情報で学習しており、日本語での学習量は英語に比べて非常に少ないと言われています。このため、日本語大規模言語モデルの開発等に当たっては、日本や日本語に関するデータが必要となり、書き言葉や話し言葉を体系的に収集した「コーパス」と呼ばれるデータベースの整備も重要となります。これまで、人間文化研究機構国立国語研究所において、約1億語の「現代日本語書き言葉均衡コーパス」や「日本語話し言葉コーパス」などの開発が行われてきましたが、大規模言語モデルの構築には、より大規模なデータが必要となります。第2節で紹介した東京工業大学と産業技術総合研究所によるSwallowの開発に当たっては、令和2年から令和5年（2020～

2023年）にかけて収集された、約634億ページに及ぶウェブのテキストから日本語のテキストを独自に抽出・精練し、約3,121億文字（約1.73億ページ）からなる日本語ウェブコーパスを構築し、学習させています²。

日本語の特性から、日本語モデルの開発には、単語の分割や品詞の特定などの言語処理を高精度に行うことも重要になります。株式会社リクルートのAI研究機関であるMegagon Labsは、人間文化研究機構国立国語研究所との共同研究成果を用いて、全世界のエンジニアやデータサイエンティストが目的に応じて日本語の自然言語処理技術を他の言語とシームレスに利用することを可能とする、日本語自然言語処理ライブラリ「GINZA」³を平成31年（2019年）4月から公開しています⁴。

また、総務省では、我が国における大規模言語モデルの開発力強化に向け、令和6年度（2024年度）から、情報通信研究機構において大量・高品質で安全性の高い日本語を中心とする学習用言語データを整備・拡充し、我が国の大規模言語モデル開発者等にアクセスを提供していくこととしています。

第4節

AIの安全性の確保に関する対策や研究開発

AI技術については、技術の精度や性能の向上だけでなく、どのようなアルゴリズムに基づいて回答しているかなどの「透明性」や、誤った回答をしていないかなどの「信頼性」等の懸念、さらには倫理的・法的・社会的課題（ELSI⁵）に対応していくことも重要です。

深層学習の結果は、多層ニューラルネット

ワーク中のリンクの重みになるので、その意味を人間が直感的に理解することは困難であり、またなぜそう判定したのか、人間に理解可能な形で理由を説明してくれないことから、ブラックボックス性が指摘されています。つまり、動作保証ができず、事故が発生しても原因解明や責任判断が不能な状況に陥りかねないリスク

1 内閣府 科学技術・イノベーション推進事務局（令和5年）「AI学習データの提供促進に向けたアクションプランver1.0」
https://www8.cao.go.jp/cstp/ai/ai_senryaku/6kai/2aidata.pdf



2 TokyoTech-LLM「Swallow Corpus」<https://tokyotech-llm.github.io/swallow-corpus>

3 Megagon Labs「GINZA：日本語自然言語処理オープンソースライブラリ」
<https://megagon.ai/jp/projects/ginza-install-a-japanese-nlp-library-in-one-step/>

4 株式会社リクルート「リクルートのAI研究機関、国立国語研究所との共同研究成果を用いた日本語の自然言語処理ライブラリ『GINZA』を公開」https://www.recruit.co.jp/newsroom/2019/0402_18331.html

5 Ethical, Legal and Social Issues

があります。

また、大規模言語モデルには、存在しない情報を、あたかも本当に存在するかのように作り出してしまい、生成された誤った情報が人間や専門家にも本物かどうか区別がつかないほど正確に見えてしまう「幻覚(ハルシネーション)」と呼ばれる問題があることが指摘されています¹。機械学習では、学習データに含まれていない未知の事例に対して、学習データから法則やルールを獲得し対応する能力である「汎化」が目指されていますが、「汎化」が適切になされなかった場合には、誤った関係性や情報を導き出すリスクがあります。

さらに、機械学習の判定結果は学習データの傾向を反映するため、仮に学習データに価値観や偏見を持ったデータが含まれていたり、データの分布に偏りがあったりすると、判定結果にもそれが反映されるリスクがあることも指摘されています。

こうした背景を踏まえ、AI戦略会議では、懸念されるリスクの具体例として、①機密情報の漏えいや個人情報の不適正な利用のリスク、②犯罪の巧妙化・容易化につながるリスク、③偽情報等が社会を不安定化・混乱させるリスク、④サイバー攻撃が巧妙化するリスク、⑤教育現場における生成AIの扱い、⑥著作権侵害のリスク、⑦AIによって失業者が増えるリスクが指摘されました²。

さらに、AIのガバナンスの在り方については、各国での検討のほか、多国間でも議論がなされています。特に生成AIの技術的な進展を踏まえた国際的なガバナンスの検討を行うため、令和5年(2023年)5月に立ち上げられた「広島AIプロセス」において議論が深められ、12月には主要7か国(G7)首脳声明で「広島AIプロセス包括的政策枠組み」等の議論の成果がG7首脳により承認されました(第3章第

2節参照)。また、米国や英国と同様に、我が国においても、令和6年(2024年)2月にAIの安全性に関する評価手法や基準の確立を目指して「AIセーフティ・インスティテュート³」を設立するとともに、総務省及び経済産業省は、AIに関係する全ての事業者を対象とした、「AI事業者ガイドライン」を令和6年(2024年)4月に取りまとめ、AIのもたらしリスクを認識しながらAIの利活用を推進していくこととしています。

また、総務省では、インターネット上の偽・誤情報の流通・拡散等に対し、令和5年(2023年)11月に「デジタル空間における情報流通の健全性確保の在り方に関する検討会」を立ち上げ、令和6年(2024年)夏頃の取りまとめに向け、生成AI・ディープフェイク技術の進展に伴うリスクへの対応の在り方を含め、総合的な対策の検討を進めています。

このようなAIガバナンスに関する取組とともに、AIの透明性・信頼性の確保を支える技術開発も進められています。例えば、外部情報の検索を組み合わせる技術である「検索拡張生成(RAG⁴)」などを活用することで、出力結果の根拠が明確になり、事実に基づかない情報の生成を抑制することが期待されています。また、情報・システム研究機構国立情報学研究所では、九州大学との連携により、画像識別AIの誤識別リスクを効果的・効率的に低減する技術を開発しました⁵。また、令和6年(2024年)4月に、同研究所内に「LLM研究開発センター」を設置し、生成AIの透明性・信頼性の確保に向けた研究開発に取り組んでいくこととしています。

また、AIが学習するデータに個人情報が含まれるケースにおいて、プライバシーやデータの利用に関する倫理的な懸念が指摘されており、プライバシーを強化、保護しながらデータ

1 岡野原大輔(2023)「大規模言語モデルは新たな知能か」岩波科学ライブラリー

2 前掲 AI戦略会議(令和5年)「AIに関する暫定的な論点整理」

3 AIセーフティ・インスティテュート <https://aisi.go.jp/>

4 Retrieval-Augmented Generation

5 情報・システム研究機構国立情報学研究所「画像識別AIの誤識別リスクを効果的・効率的に低減する技術を開発～自動運転システムにおける安全性ベンチマークにて効果を検証～」<https://www.nii.ac.jp/news/release/2023/0317.html>

分析を可能にする技術も開発されてきています。具体的には、分析対象データ保護のためのデータ匿名化技術、データベース問合せ結果保護のための差分プライバシー技術、計算過程におけるデータ内容保護のための秘密計算(若しくは秘匿計算)技術などについて、大学や企業等で研究開発が進められており、例えば、NTTが開発した秘密計算技術が、2024年3月、ISO国際標準に採択¹されました。これらの技術のAI分野への適用が進められています。

さらに、セキュアなAIシステムの構築に向けて、内閣府科学技術・イノベーション推進事務局及び内閣サイバーセキュリティセンターは、令和5年(2023年)11月に、英国国家サイバーセキュリティセンターが米国サイバーセキュリティ・インフラストラクチャー安全保障

庁等と共に作成した「セキュアAIシステム開発ガイドライン」の「共同署名」に加わるとともに、令和6年(2024年)1月に、豪州サイバーセキュリティセンターがカナダ、ニュージーランド、英国、米国の関係当局とともに作成した「AI使用に関する国際ガイダンス」の「共同署名」に加わり、文書を公開しました^{2,3}。

また、経済安全保障推進会議及び統合イノベーション戦略推進会議の下、内閣府、文部科学省及び経済産業省が中心となって、府省横断的に実施している「経済安全保障重要技術育成プログラム⁴」において、「人工知能(AI)が浸透するデータ駆動型の経済社会に必要なAIセキュリティ技術の確立」についても、研究開発構想の一つとして掲げられ、令和5年度(2023年度)に公募が行われました。

第5節 人材育成

最先端のAIの知識・技能を持つ人材育成の取組も進められています。文部科学省では、AI分野及びAI分野における新興・融合領域(次世代AI分野)の人材育成及び先端的研究開発を推進するため、「国家戦略分野の若手研究者及び博士後期課程学生の育成事業(BOOST)次世代AI人材育成プログラム⁵」を創設し、令和6年(2024年)1月に科学技術振興機構で博士後期課程の学生に対し公募を開始しました。同分野に資する研究開発に取り組みようとする博士後期課程学生に対して、十分な

生活費相当額及び研究費を支援することで、当該国家戦略分野の研究者層を厚くし、イノベーション創出や産業競争力を強化することを目的としています。

また経済産業省では「未踏事業」において、AIに限らずITを駆使してイノベーションを創出することのできる独創的なアイデアと技術を有するとともに、これらを活用する優れた能力を持つ、突出した人材の発掘・育成を進めています(第1-2-9表)。

1 日本電信電話株式会社「NTTの秘密計算技術がISO国際標準に採択～データ利活用とプライバシー保護を両立する高速な秘密計算技術を実現～」<https://group.ntt.jp/newsrelease/2024/03/21/240321b.html>

2 内閣府(2023)「セキュアAIシステム開発ガイドラインについて」<https://www8.cao.go.jp/cstp/stmain/20231128ai.html>

3 内閣府(2024)「AI使用に関する国際ガイダンスへの共同署名について」<https://www8.cao.go.jp/cstp/stmain/20240124.html>

4 内閣府「経済安全保障重要技術育成プログラム」https://www8.cao.go.jp/cstp/anzen_anshin/kprogram.html



5 科学技術振興機構「国家戦略分野の若手研究者及び博士後期課程学生の育成事業(BOOST)次世代AI人材育成プログラム」<https://www.jst.go.jp/jisedai/boost-s/index.html>



■ 第1-2-9表 / AI人材育成事業の例

事業名	対象	概要
数理・データサイエンス・AI教育プログラム認定制度(文部科学省)	大学及び高等専門学校	数理・データサイエンス・AIに関する知識及び技術について体系的な教育を行うプログラムを文部科学大臣が認定及び選定して奨励する。 令和5年(2023年)8月時点の総認定数はリテラシーレベルが382件、応用基礎レベルが147件である。
K-DASH (KOSEN Mathematics, Data science and AI Smart Higher Educational Community) ¹ (文部科学省、国立高等専門学校機構)	高等専門学校	旭川高専と富山高専を拠点校とし、全国の国立高専と連携し数理・データサイエンス・AI教育を実施。『AI×DS(データサイエンス)で加速する高専生』『AIとデータでスペシャリストへ加速する高専生』を全国高専で育成すべき人材像に掲げ、卒業生の100%にリテラシーレベル、20%以上に専門レベルを習得させ、1%のトップレベルから研究者、起業家などの人材を輩出することを目指している。
国家戦略分野の若手研究者及び博士後期課程学生の育成事業(BOOST)次世代AI人材育成プログラム(文部科学省、科学技術振興機構)	若手研究者及び博士後期課程学生	次世代AI分野(AI分野及びAI分野における新興・融合領域)に資する研究開発に取り組もうとする若手研究者及び博士後期課程学生に対して支援を行う。
ACT-X(文部科学省、科学技術振興機構)	博士の学位取得後8年未満の若手研究者(大学院生を含む。)	独創的・挑戦的なアイデアを持つ若手研究者を支援する本プログラムの中で、AIに関するものとして「数理・情報のフロンティア」、「AI活用で挑む学問の革新と創成」、「次世代AIを築く数理・情報科学の革新」の領域を設け、令和5年度(2023年度)末時点で計187人の若手研究者を支援している(終了した課題を含む)。
未踏事業(経済産業省、情報処理推進機構)	未踏IT人材発掘・育成事業: 25歳未満 未踏アドバンスト事業: 条件なし 未踏ターゲット事業: 条件なし	AIに限らずITを駆使してイノベーションを創出することのできる独創的なアイデアと技術を有するとともに、これらを活用する優れた能力を持つ、突出した人材を発掘・育成することを目的とし、若い人材の発掘・育成を行う「未踏IT人材発掘・育成事業」、ITを活用した革新的なアイデア等を有し、ビジネスや社会課題の解決につなげる人材を育成する「未踏アドバンスト事業」、次世代ITを活用して世の中を抜本的に変えていける先進分野の人材を育成する「未踏ターゲット事業」を実施している。

資料：文部科学省作成

¹ 国立高等専門学校機構「高専発!『Society 5.0型未来技術人材』育成事業(GEAR 5.0/COMPASS 5.0)」
<https://www.kosen-k.go.jp/about/profile/gear5.0-compass5.0.html>

岐阜大学では、最先端AI人材の育成を目指しており、「数理・データサイエンス・AI教育プログラム認定制度」に認定されたほか、実践的なAI教育としてNVIDIA社の「Jetson Nano」を工学部の実験科目に大規模導入しました。Jetson Nanoは小型コンピュータでありながらAIを高速で実行できるエッジ端末です。当該科目では、エッジAIの仕組みを理解した上で学生自身がAIプロジェクトを設計、実装、評価するアクティブラーニングを促します。AIによる画像認識、ロボット制御、音声認識技術について理解し、それらをJetson Nano上で組み込みAIシステムの開発を行います。成果物をプレゼンテーション、デモンストラレーションし、最終的にAIとJetsonの基礎的な実践スキルを習得した証明であるNVIDIA社の認定資格、Jetson AI Specialistの取得を目指しています（第1-2-10図）。

富山高等専門学校（富山高専）は、地域企業と連携したDXやAI・データ活用状況の調査、報告書をまとめる取組を実施し、「数理・データサイエンス・AI教育プログラム認定制度リテラシープラス」に高専では2校目として選定されています（第1-2-11図）。

また、富山高専の特徴的な取組の一つであるAI教育に外部人材を登用する「AI副業先生」では、株式会社ビズリーチと国立高等専門学校機構との連携により、ビジネスの現場でAIを活用して働く民間人材を教員として採用し、社会変化やニーズに応えられるAI技術を教えています。

香川県三豊市及び東京大学大学院工学系研究科松尾豊教授が平成31年（2019年）4月に設立した一般社団法人みとよAI社会推進機構（MAiZM）では、香川高等専門学校と連携し、AIサマースクールの実施や高専発スタートアップの支援など、人材育成や地域社会の課題解決に向けた取組が行われています。

■第1-2-10図／ツールキットJetson Nanoを使用している様子



提供：東海国立大学機構岐阜大学

■第1-2-11図／DXやAI・データ活用状況に関する地域企業へのオンライン調査の様子



提供：富山高等専門学校

本章では、我が国でのAI開発の事例を見ってきましたが、我が国の強みの一つは、自動車、エレクトロニクス、ロボティクスなどの分野での長年の経験と技術の蓄積です。このような産業基盤を生かしながら、大学や研究機関による高い技術力と研究開発能力により、AIの更なる開発とともに、多くの技術分野でのAIの活用や第4章で紹介する科学分野での活用によるイノベーションの創出が期待されています。