安全なデータ連携による最適化ＡＩ技術の研究開発

基本計画書

１．目　的

我が国では、統合イノベーション戦略推進会議が中心となって、Society 5.0 の実現を通じて世界規模の課題の解決に貢献するとともに、我が国自身の社会課題の克服や産業競争力の向上に向けたＡＩに関する総合的な政策パッケージとして「ＡＩ戦略」が取りまとめられており、令和４年４月には、その最新版として「ＡＩ戦略 2022」が策定されたところである。

「ＡＩ戦略 2022」の中では、ＡＩの社会実装を更に推進するため、５つの戦略目標が掲げられており、そのうちの１つとして、「我が国が、実世界産業におけるＡＩの応用でトップランナーとなり、産業競争力の強化が実現されること」が挙げられ（戦略目標２）、「実世界産業」領域に含まれる系統的に取得されていない膨大な情報をＡＩに活用することで、我が国産業の労働生産性の向上等による産業競争力の向上や、SDGs 達成への貢献等を目指すこととしている。

この戦略目標を達成する上での具体的な取組としては、ＡＩの適用領域の拡大や、我が国産業が有する分野毎の高品位データの利活用や他分野との連携等の、ＡＩの利活用を支えるデータの充実に取り組むとともに、日本が強みを有する分野とＡＩの融合が求められている。これにより、我が国が強みを有する産業の競争力を向上させるだけでなく、他分野と連携することで、我が国全体の競争力の向上を目指すものであるが、一般に、そのような産業界が有するデータ群は貴重なものであり、いわゆる社外秘として容易には外部に提供され得ないものであることが、これまで他分野での活用を阻害する要因の１つとなっていた。

このため、同戦略では、データ利活用のための必要な要素技術の１つとして「プライバシーや機密情報を保護しながら学習可能な連合学習（Federated Learning）など一連の技術の一層の研究開発・社会実装の推進」を挙げている。

このような背景を踏まえ、本研究開発を実施する前提として、我が国が先端的ＡＩ技術の研究開発において世界をリードし、多様な者が有するデータを安全に連携させることを可能とする「データ連携ＡＩプラットフォーム」を世界に先駆けて創出し、分野横断的なＡＩの応用に関して我が国がトップランナーになることを目的とする。当該目的の下、本研究開発では、データを共有せずにＡＩ学習に活用する連合学習技術と、様々な種類のデータを組み合わせてＡＩ学習に活用するマルチモーダルＡＩ技術及びエッジ環境でＡＩ学習を行うエッジＡＩ技術とを組み合わせることにより、プライバシーデータや機密データ等を含め、実空間に存在する多様なデータを安全に連携させ分野横断的な課題解決を可能とする分散型機械学習技術を確立することを目指す。

２．政策的位置付け

「新しい資本主義のグランドデザイン及び実行計画」（令和４年６月７日閣議決定）において、「ＡＩ技術を基にした実践・試行錯誤の蓄積が重要であり、ディープラーニングを重要分野として位置付け、企業による具体的ニーズを念頭に置き、その実装・開発を推進する」、「データをできるだけ多く利用できる環境を整えるべく、プライバシー等の理由により秘匿化された情報について、秘匿化したままで機械学習の処理を行うことができるよう、技術開発を推進する」とされている。

「デジタル社会の実現に向けた重点計画」（令和４年６月７日閣議決定）において、「今後の更なるＡＩの実用化に向けて、「ＡＩ戦略 2022」 に基づき、ディープラーニングを重要分野として位置付け、企業による実装を念頭において取り組む」、「ＡＩ利活用を支えるデータの充実等に向けて、データの秘匿性を担保したままで機械学習の処理等を行うための研究開発、研究データ基盤の改善などのほか、データの取扱いルールについての再点検その他の環境整備に取り組む」とされている。

「統合イノベーション戦略 2022」（令和４年６月３日閣議決定）において、「ＡＩの社会実装の更なる推進のため、画像認識、自然言語処理等での広範かつ効果的な活用が期待されるディープラーニングを重要分野として位置付け、企業による実装を念頭に置きつつ、ＡＩの信頼性向上、ＡＩ利活用を支えるデータの充実、ＡＩを巡る人材や技術情報、データ取扱いルール等の追加的な環境整備、政府におけるＡＩ利活用の推進、我が国が強みを有する分野とＡＩとの融合に力点を置いて取り組む」とされている。また、ＡＩ技術に係る今後の取組方針として、「エッジ環境のＩｏＴデータを共有せず実空間の分野横断的な行動リスク予測を可能にする分散連合型のマルチモーダル・クロスモーダルＡＩ技術の研究開発」が挙げられている。

「ＡＩ戦略 2022」（令和４年４月 22 日統合イノベーション戦略推進会議決定）において、「説明可能なＡＩ(Explainable AI, XAI)やプライバシーや機密情報を保護しながら学習可能な連合学習(Federated Learning)など一連の技術の一層の研究開発・社会実装の推進とプラットフォーム化、およびその運用におけるリーダーシップが重要となる」とされている。また、「「説明可能なＡＩ」（Explainable AI）など「責任あるＡＩ」（Responsible AI）の実現」や「秘匿データの効果的な利用につながる、サイバーセキュリティとＡＩの融合領域の技術開発等の推進」という目標を達成するための取組として、「エッジ環境の IoT データを共有せず実空間の分野横断的な行動リスク予測を可能にする分散連合型のマルチモーダル・クロスモーダル AI 技術の研究開発」が挙げられている。

「Beyond 5G に向けた情報通信技術戦略の在り方 ー強靭で活力のある 2030 年代の社会を目指してー 中間答申」(令和４年６月 30 日情報通信審議会中間答申)において、Beyond 5G に向けて産学官全体で取り組むべき研究開発課題の１つとして「Beyond 5G サービス・アプリケーション技術」が挙げられており、その主な要素技術として「連合機械学習 (Federated Learning)」が挙げられている。

３．目　標

（１）政策目標（アウトカム目標）

ＡＩによる我が国の社会課題の解決や産業競争力の向上を実現していくためには、我が国が強みを有する分野を含め、分野毎に有するデータを連携させデータを分野横断的に活用することにより、ＡＩの高性能化を図り、より複雑な課題をＡＩにより解決していく必要がある。

しかし、異なる者（民間企業、地方公共団体等）の間でデータを連携させることを考えた場合、各者で有するデータを相互に利用することはＡＩ学習にとって非常に有効である一方、プライバシーデータや機密データ等の取扱いに注意を要するデータを他者にそのまま共有することが困難であり、現実には、データの連携が進まないという問題が存在している。この問題を解決する手段として、データそのものを他者に共有せずともＡＩ学習に活用することを可能とする技術、更には、多様な者のコミュニティの場となり、当該技術によって多様な者が有するデータを安全に連携させることを可能とする「データ連携ＡＩプラットフォーム」が求められる。

そこで、本研究開発では、（２）のとおり、３つの要素技術を確立し、更にこれらを組み合わせることで、実空間に存在する多様なデータを安全に連携し分野横断的な課題解決を可能とする分散型機械学習技術を確立することを研究開発目標とし、本研究開発終了後、確立した技術を活用した「データ連携ＡＩプラットフォーム」が創出され、分野を横断したデータ活用が進められることにより、我が国の社会課題の解決や我が国産業の労働生産性の向上等による産業競争力の向上に貢献することを政策目標（アウトカム目標）とする。

（２）研究開発目標（アウトプット目標）

本研究開発では、要素技術として、ＡＩ学習用に実空間から収集するデータの量や粒度の差異を吸収しつつ、多様なデータを組み合わせ複雑な予測を可能とするロバストなマルチモーダルＡＩ技術、エッジ環境の限られた計算資源の規模に応じて効率的に学習を行うエッジＡＩ技術、多数のエッジ環境間におけるデータの偏りを前提とした高精度な連合学習技術を確立し、更にこれらを組み合わせることで、実空間に存在する多様なデータを安全に連携し分野横断的な課題解決を可能とする分散型機械学習技術を確立することを研究開発目標（アウトプット目標）とする。

４．研究開発内容

（１）概要

（２）に記載する３つの要素技術の研究開発を行う。更に、それらを組み合わせた分散型機械学習システムを試作し、具体的な社会実装シーンを想定した技術実証を行う。

（２）技術課題

ア）マルチモーダルＡＩ技術

従来のＡＩ学習では、単一の種類のデータ（例えば、テキストデータのみ）が学習に用いられる（一般に「シングルモーダルＡＩ」と呼ばれる。）。これに対し、IoT の進展も相まって、実空間に存在する様々な IoT デバイスから収集される IoT データを含め、複数種類のデータ（例えば、テキストデータと画像データ）を組み合わせて学習に用いることで、より複雑な予測を可能とする「マルチモーダルＡＩ技術」が近年登場している。

一方、複数種類のデータの収集に当たっては、特に IoT データにおいて、センサ数やセンシング頻度、都市部と地方部における人口差、平時と異常発生時の期間差、プライバシーデータの活用可否等に起因し、収集された場所や期間等によってデータの量や粒度が異なる等の収集データの差異が発生し、マルチモーダルＡＩの予測性能を著しく低下させることが課題となっている。

特に、多数のエッジ環境において個別に収集したデータを基にそれぞれでＡＩ学習を行う連合学習を想定した場合、エッジ環境間で収集データに差異が生じ、十分なデータを収集できないエッジ環境では学習による予測性能が著しく低下する事態が想定される。連合学習は、各エッジ環境で学習したＡＩモデルをクラウド環境に集約してクラウド環境上のＡＩモデルの学習を行い、その学習結果を各エッジ環境に還元する仕組みとなっており、各エッジ環境での学習による予測性能の差を補完する技術ではあるが、全体としてより高精度な予測性能を達成するためには、十分なデータを収集できないエッジ環境を含め、各エッジ環境において予測性能を向上させることが求められる。

そこで本研究開発においては、多様なデータを組み合わせ複雑な予測を可能とする大規模マルチモーダル深層学習モデルの構築技術、及び、実空間から収集するデータの差異を吸収可能な、ロバストなマルチモーダルＡＩ技術についての研究開発を行う。

その実施に当たっては、技術課題「イ）エッジＡＩ技術」及び技術課題「ウ）連合学習技術」と連携して、３つの要素技術を組み合わせて分散型機械学習システムを試作し、具体的な社会実装シーンを想定した技術実証を行うことにより、その有用性について検証するとともに、検証結果を踏まえて各要素技術の更なる改善を行う。

イ）エッジＡＩ技術

一般に、通常のＡＩ学習においては、大量の学習用データをクラウド環境等に集約し、豊富な計算資源を活用してＡＩ学習を行う（一般に「データ集中型ＡＩ」等と呼ばれる。）。これに対し、よりデータの発生源・保管場所に近いエッジ環境において、限定的な計算資源を活用して、比較的少量のデータによりＡＩ学習を行う技術をエッジＡＩ技術という。

エッジ環境で学習を行うため、データを手元に保存したままで学習を行うことが可能であり、クラウド環境へデータを集約するための通信コストを抑えること

等が可能となる一方で、１つの建物内や１つの地区内等の、小規模な範囲をカバーするエッジ環境の限定的な計算資源では、大規模な深層学習モデルの学習が難しいことが課題となっている。

そこで本研究開発においては、マルチモーダル深層学習モデルを対象に、エッジ環境の限られた計算資源の規模に応じて、効率的に学習を行う技術の研究開発を行う。

その実施に当たっては、技術課題ア）同様、３つの要素技術を組み合わせて分散型機械学習システムを試作し、具体的な社会実装シーンを想定した技術実証を行うことにより、その有用性について検証するとともに、検証結果を踏まえて各要素技術の更なる改善を行う。

ウ）連合学習技術

従来のＡＩ技術では、学習用データをクラウド環境等に集約してＡＩ学習を行う。この場合、クラウド環境等への集約が難しいプライバシーデータや機密データ等の取扱いに注意を要するデータをＡＩ学習に活用することが困難という課題があり、この課題解決のための技術的アプローチが模索されている。

連合学習技術は、そのアプローチの１つであり、データそのものを集約するのではなく、データはエッジ環境でのＡＩ学習にのみ用いられ、クラウド環境でのＡＩ学習には、エッジ環境で学習を行ったＡＩモデルを集約し活用する技術である。すなわち、本技術により、データを利用者側のエッジ環境に留め置き秘匿性を保持しつつも、データの代わりにデータの傾向のみを反映したＡＩモデルをクラウド環境に集約することにより、データを直接的に連携させずとも、ＡＩモデルを介した間接的な手法により安全に連携させることが可能となる[1]。

ただし、取扱いに注意を要するデータをＡＩ学習に活用しやすくなる一方、個々のエッジ環境で収集されるデータの量や粒度が異なる等の偏りがあるデータで学習したＡＩモデルを単純にクラウド環境に集約し学習しただけでは、ＡＩモデルの予測性能が劣化してしまうという課題がある(特にマルチモーダルＡＩは、多様な IoT データ等を活用するが故に、データの偏りが発生しやすい。)。

そこで本研究開発においては、マルチモーダル深層学習モデルを対象に、多数のエッジ環境間におけるデータの偏りを前提とした高精度な連合学習技術の研究開発を行う。

その実施に当たっては、技術課題ア）及びイ）同様、３つの要素技術を組み合わせて分散型機械学習システムを試作し、具体的な社会実装シーンを想定した技術実証を行うことにより、その有用性について検証するとともに、検証結果を踏まえて各要素技術の更なる改善を行う。

[1] ただし、ＡＩモデルを分析することによりエッジ環境のデータを推測する攻撃等、連合学習技術に対する脅威も存在することに留意する必要がある。実際に、当該攻撃への対策に係るセキュリティ技術等についても研究が進められている。

（３）到達目標

前提条件として、３種類以上[2]のマルチモーダルデータを組み合わせ複雑な予測を可能とする１億 5000 万パラメータ規模[3]のマルチモーダル深層学習モデルを想定し、実空間から収集するデータの差異、エッジ環境の計算資源の規模、及び、エッジ環境間におけるデータの偏りを考慮して、高精度に、効率的に、かつ、安全にデータを連携させ、ＡＩ学習を行う分散型機械学習技術を確立することを本研究開発全体の到達目標とし、その実現のための各要素技術の到達目標を以下のとおり設定する。

ア）マルチモーダルＡＩ技術

３種類以上のマルチモーダルデータを組み合わせ複雑な予測を可能とする１億 5000 万パラメータ規模のマルチモーダル深層学習モデルの構築技術の確立、及び、収集された場所や期間等によってデータの量や粒度が 1/10 程度[4]の少量になってしまう場合（あるエッジ環境にて収集されるデータの量や粒度が、他のエッジ環境の 1/10 程度の少量になってしまう場合）でも、予測性能を概ね維持できるようなロバストなマルチモーダルＡＩ技術の確立を目標とする。

イ）エッジＡＩ技術

マルチモーダル深層学習モデル[5]を対象に、１つの建物内や１つの地区内等の、小規模な範囲をカバーするエッジ環境の限られた計算資源の規模に応じて、効率的に学習を行う技術の確立を目標とする。

ウ）連合学習技術

マルチモーダル深層学習モデル[5]を対象に、数百程度[6]のエッジ環境間におけるデータの偏りを前提に、データ集中型ＡＩに近い予測性能を実現可能な連合学習技術の確立を目標とする。

---

[2] 少量データからＡＩ学習を行う技術について、画像データとテキストデータの２種類のマルチモーダルデータを対象に研究が行われている例がある。将来的には、より多種類のマルチモーダルデータを対象とした、少量データから学習可能なマルチモーダルＡＩ技術が求められると想定されることから、３種類以上のマルチモーダルデータを取り扱うことを前提条件として設定。

[3] 画像データを取り扱う深層学習モデルとして、画像分類の分野で最も広く活用されているモデルの１つである EfficientNet のパラメータ数は約 6600 万（EfficientNet-B7）であり、画像データに加えテキストデータや数値データも取り扱うマルチモーダル深層学習モデルのパラメータ数としては、１億超規模が妥当であると考えられる。近年、深層学習モデルはますます大規模化している傾向にあることを踏まえ、研究開発終了時期には、よりパラメータ数の多い１億 5000 万パラメータ規模のマルチモーダル深層学習モデルが主流になると想定し、当該規模を前提条件として設定。

[4] センサ数やセンシング頻度、都市部と地方部における人口差、平時と異常発生時の期間差、プライバシーデータの活用可否等に起因し、データの量や粒度が、オーダーが異なる程に少なくなってしまう場合でも、予測性能を概ね維持することを目標に、少量データの基準を 1/10 程度と設定。

[5] ア）で構築されるマルチモーダル深層学習モデルによっては、エッジ環境ではモデルの全体ではなく一部分のみを学習する場合も想定されることから、イ）及びウ）においては、マルチモーダル深層学習モデルのパラメータ規模は到達目標として設定しない。

[6] 研究開発終了時期には、エッジ環境の高密度化が進み、１つのクラウド環境（データセンター）に対し数百程度のエッジ環境が配置されることが想定されていることから、数百程度のエッジ環境を前提とした目標を設定。

５．研究開発期間

初回契約締結日から令和７年度

６．その他　特記事項

（１）特記事項

提案者は、次の課題ア）、課題イ）、課題ウ）のいずれか又は複数の課題に提案することができる。なお、いずれの研究開発の受託者も相互に連携、協力して研究開発を行う。

また、原則、課題ア）の受託者が本研究開発課題全体のとりまとめを行うものとする。ただし、別の課題（課題イ）又は課題ウ））の受託者が本研究開発課題全体のとりまとめを行うとする提案も可とする。

ア）マルチモーダルＡＩ技術
イ）エッジＡＩ技術
ウ）連合学習技術

（２）提案及び研究開発に当たっての留意点

① 提案に当たっては、国際的な研究開発の状況を踏まえた上で、本基本計画書に記されているアウトプット目標に対する達成度を評価することが可能な具体的な評価項目を設定し、各評価項目に対して可能な限り数値目標を定めること。

なお、本基本計画書に記されているアウトプット目標については、現時点における技術トレンドを踏まえて設定したものである。一方、研究開発及び社会実装において世界のトップランナーを目指すという本研究開発の趣旨を踏まえれば、本研究開発の実施に際しては、国際的な研究開発の状況・技術トレンドを踏まえた上で、目標を更新していく必要がある。このため、研究開発期間中に必要に応じて、数値目標の上方修正等の更新について、総務省と協議を行いながら研究開発を進めること。

② 提案に当たっては、研究開発成果の最大化のため、研究開発課題として設定している３つの要素技術に加えて、深層学習に限らないＡＩ技術、安全なデータ連携を担保するセキュリティ技術等、その他の技術を組み合わせ、新規性・独創性の高い研究開発を提案すること。

③ アウトプット目標を達成するための研究方法、実用的な成果を導出するための共同研究体制又は研究協力体制、及び達成度を客観的に評価するための実験方法について、具体的に提案書に記載すること。特に、本研究開発の実施に当たっては、その前提として、マルチモーダル深層学習モデルの学習や評価のために必要なデータの

整備を迅速に行う必要があるが、当該データの整備に係る計画（提案者が既に有するデータの活用に関する計画を含む）についても、具体的に提案書に記載すること。

④ 複数機関による共同研究を提案する際には、研究開発全体を整合的かつ一体的に行えるよう参加機関の役割分担及び分担する技術間の連携を明確にし、インターフェースを確保するとともに、研究開発期間を通じて継続的に連携するための方法について具体的に提案書に記載すること。

⑤ 研究開発の実施に当たっては、関連する要素技術間の調整、成果の取りまとめ方等、研究開発全体の方針について幅広い観点から助言を頂くとともに、実際の研究開発の進め方について適宜指導を頂くため、学識経験者、有識者等を含んだ研究開発運営委員会等を開催する等、外部の学識経験者、有識者等を参画させること。

⑥ アウトカム目標の達成に向けた適切な研究成果（アウトプット等）の取扱方策（研究開発課題の分野の特性を踏まえたオープン・クローズ戦略を含む。）について提案すること。

⑦ 本研究開発成果を確実に展開し、アウトカム目標を達成するため、事業化目標年度、事業化に至るまでの実効的な取組計画（事業化に向けた具体的な計画、体制、資金等）についても具体的に提案書に記載すること。

⑧ 本研究開発は総務省施策の一環として取り組むものであることから、総務省が受託者に対して指示する、研究開発に関する情報及び研究開発成果の開示、関係研究開発プロジェクトとのミーティングへの出席、シンポジウム等での研究発表、共同実証への参加等に可能な限り応じること。

（3）人材の確保・育成への配慮

① 研究開発によって十分な成果が創出されるためには、優れた人材の確保が必要である。このため、本研究開発の実施に際し、人事、施設、予算等のあらゆる面で、優れた人材が確保される環境整備に関して具体的に提案書に記載すること。

② 若手の人材育成の観点から行う部外研究員受け入れや招へい制度、インターンシップ制度等による人員の活用を推奨する。また、可能な限り本研究開発の概要を学会誌の解説論文で公表する等の将来の人材育成に向けた啓発活動についても十分に配慮すること。これらの取組予定の有無や計画について提案書において提案すること。

（4）研究開発成果の情報発信

① 本研究開発で確立した技術の普及啓発活動を実施するとともに、その活動計画・方策については具体的に提案書に記載すること。

② 研究開発成果については、原則として、総務省としてインターネット等により発信を行うとともに、マスコミを通じた研究開発成果の発表、講演会での発表等により、広く一般国民へ研究開発成果を分かりやすく伝える予定である。このため、研究成果に関する分かりやすい説明資料や図表等の素材、英訳文書等を作成し、研究成果報告書の一部として報告する旨を提案書に記載すること。更に、総務省が別途指定する成果発表会等の場において研究開発の進捗状況や成果について説明等を行う旨を提案書に記載すること。

③ 本研究開発終了後に成果を論文発表、プレス発表、製品化、Web サイト掲載等を行う際には「本技術は、総務省の「安全なデータ連携による最適化ＡＩ技術の研究開発」において、委託を受けて実施した研究開発による成果である。」旨の注記を発表資料等に都度付すこととする旨を提案書に明記すること。