資料２
科学技術・学術審議会 情報委員会
情報科学技術分野における戦略的重要
研究開発領域に関する検討会（第2回）
令和６年5月14日

「安全なデータ利活用のための高度な分散処理などに関する研究領域」に関して

2024年5月14日

株式会社KDDI総合研究所
小野　智弘

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

自己紹介

■ 現在

- **株式会社KDDI総合研究所　Human-Centered AI研究所　所長**
  https://researchmap.jp/chihiro_ono

■ 研究キーワード

- **実世界AI、行動変容、人口動態分析、情報推薦**

■ データ利活用に関連する取り組み

- **防災：**　GSMA Bigdata for Social Good member 2017-，第2期SIP「国家レジリエンス (防災・減災) の強化」 2018 – 2023.3
- **マーケティング：**JST CREST 人工知能「異種ドメインユーザの行動予測を可能にするペルソナモデルの転移技術」2020 ～2024.3
- **スマートシティ：**総務省受託「安全なデータ連携による最適化 AI 」2023.7～

■ 学会・社外活動

- **JSTさきがけ「文理融合による人と社会の変革基盤技術の共創」領域アドバイザー**
- **人工知能学会副会長(2023.6-)、情報処理学会理事(2023.6-)**

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

背景

■ 現在の潮流

データ駆動型社会が進みつつある一方で・・・

- **デジタル分野における国際的競争力の低下**
  - 膨大なデータと計算資源を持つメガテックの独壇場（大規模基盤モデルなど）
  - 海外プラットフォーマーによるデータ支配の脅威（スマートウォッチ/スピーカーなどでのパーソナルデータの収集など）
- **個人情報保護やデータ保護の流れ**
  - 広島AIプロセス(2023.12) 行動規範11.適切なデータインプット対策を実施し、個人データ及び知的財産を保護する。
- **企業の顧客データ活用のデータ利活用状況は不十分（アメリカの1/3以下の水準*1）**

■ 今後の見通し

- **オープンなデータは今後枯渇とも　(*2, 高品位データは2026年など）**
- **オープンなデータのみでできることは限界で、企業などが保有するドメイン特化データがより重要に**
- **企業などが保有するデータは表に出せないものも多数あり、データを全て中央に集めるという姿勢だけでは成り立たない**

→ **分散型AI＝様々な組織・デバイスに散在するデータ・計算資源を安全・効率的に活用の取り組みが重要**

*1) 独立行政法人　情報処理推進機構、DX白書2023より

*2)Will We Run Out of ML Data? Evidence From Projecting Dataset Size Trends – Epoch AI

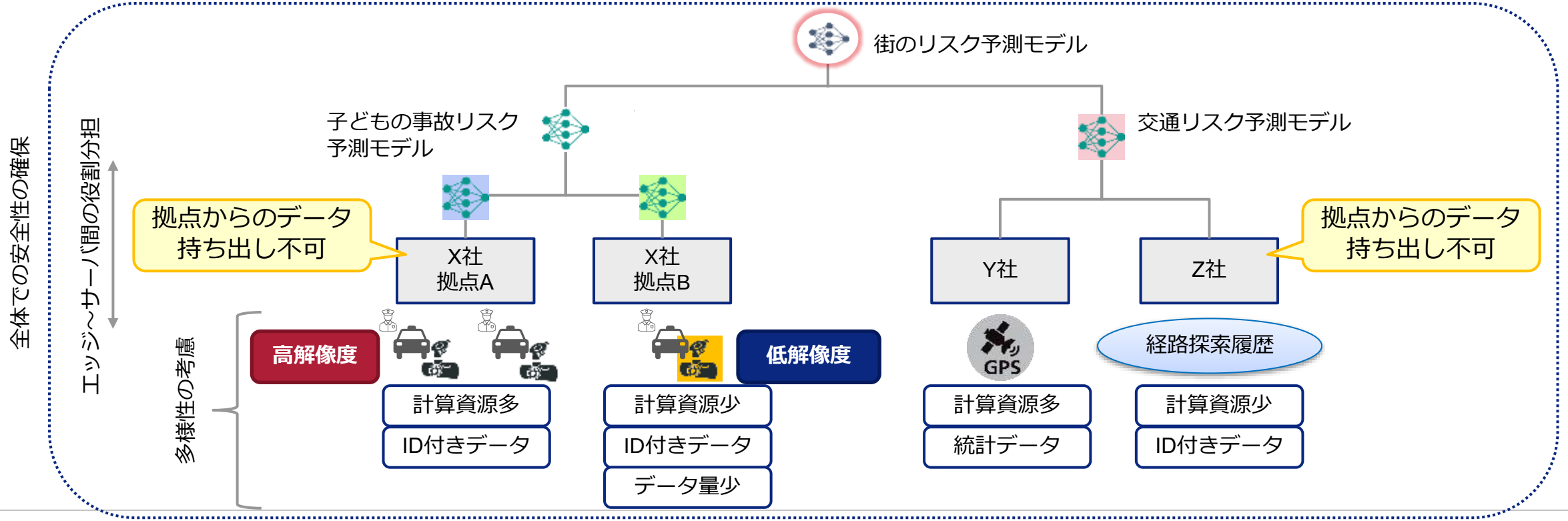
Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

背景 Cont.

■ **分散型AI(様々な組織・デバイスに散在するデータ・計算資源を安全・効率的に活用)を実現するの留意点**

1. **エッジ～サーバ間の連携：**エッジ(組織/デバイス）～サーバ間の役割分担(学習・推論)、相互フィードバックの方法
2. **安全性の確保：**分散環境全体での、データプライバシーの保護
3. **組み合わせの必要性：**組織内の複数拠点だけでなく複数の組織のデータの組み合わせ利用のニーズも高い
4. **データや環境の多様性：**組織毎・センサデバイス毎のデータの質や量が大きく異なる。計算資源も異なる
5. **企業事情の多様性：**各組織が保有するデータの利用許諾内容、求める内容（安全性重視/精度重視など）も異なる

留意点の例：複数組織のIoTデータからの各種リスク予測

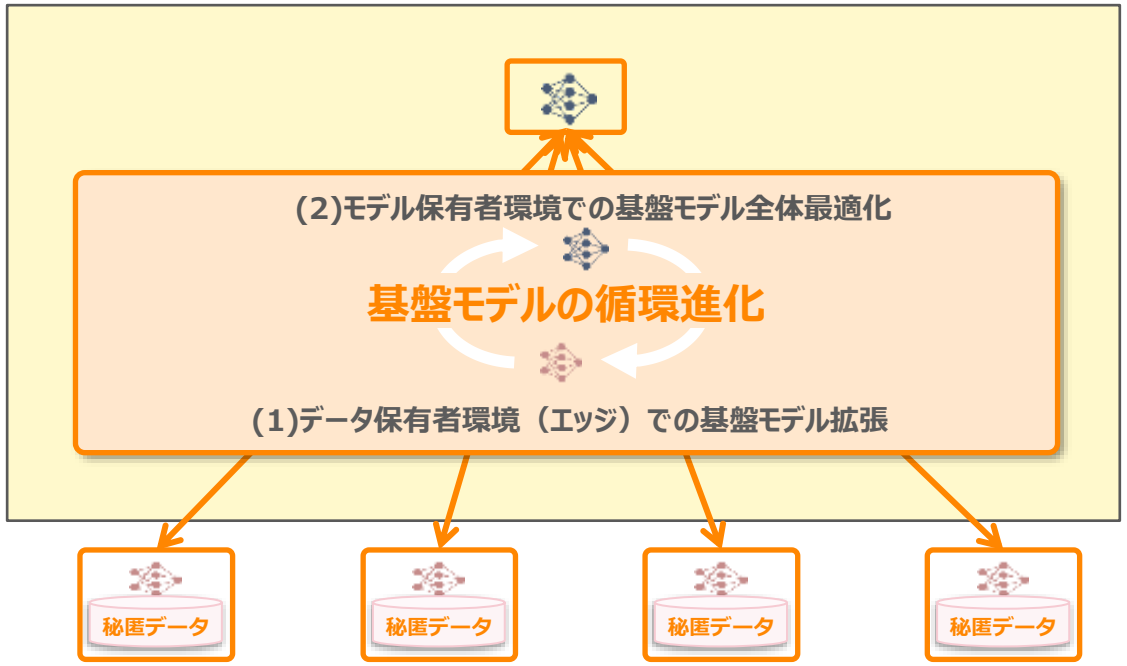


Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

AI基盤モデルの循環進化

- (1) 様々な組織やデバイスがエッジ側で収集し保有している多様・不均衡・中小規模のデータを利用して、エッジ側で基盤モデルを**個別適応化(Fine Tuning)**するとともに、(2)それらモデルの差分を基盤モデルに集約し**全体最適化**することを繰り返すことで、タスクや環境・データの変化に応じて基盤モデルを様々な利用を通じ安全かつ効率的に**循環進化**させる高度な**分散機械学習**技術を確立する

- 各組織が個別にモデルを開発するのでなく、業界団体、地域・都市などの**共通課題に特化した基盤モデル**を利用者主導で**共創する**ことで質の高いモデルを開発・運用



日本の強み

センシング、ロボット、エッジコンピューティング等を活用した、エッジ側での稠密・高品位なデータの収集とAI処理

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

期待されるユースケースの例

街の変化に伴う予測モデルの早期更新

大規模商業施設がオープンして**市民の活動が大きく変化した**ため、街中の様々な店舗や交通機関などが、多種類のデータを組み合わせて少ない期間のデータ量で早期に各種予測モデル(来店・需要・・）を更新したい

上記実現にあたり、組織毎の計算資源や技術力に差があるため、エッジ～サーバ間の役割分担（学習・推論それぞれ）を柔軟に調整したい

一人暮らしの事故検知と対応

・各家庭に置かれている見守りロボットが新たな転倒パターンを自動的に検知。家庭毎のデータを共有せずに新たなパターンに対応可能なモデルを更新して予測可能に

・日本ならではの親切な応対のために、見本となる介助者からの多様なデータを統合して応対
→丁寧な日本人介助者を手本としたモデルは海外へ販売可能

大規模災害時の即時対応

- 地震で基地局の倒壊などにより、複数エリアの**スマートホン位置情報が欠損や遅延している状況で、**収集できるデータや経路探索データなどを活用して今後の人流や交通流を予測したい

- 予測した人流・被災状況（複数ドローン撮像）・各社の在庫状況（共有不可）から救援物資を配送計画を立案したい

組織横断のマーケティング

顧客理解・リーチの拡大のため、データの提供側組織・利用側組織それぞれに求める精度、手軽さ、導入コストが異なる中で、柔軟かつ最適な突合・分析手段を提供
（データクリーンルーム(*)など）

(*)プライバシを保護しつつ各組織が保有する情報を突合・分析する環境

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

「AI基盤モデルの循環進化」に今後重要となる取り組み

■ **今後はデータ・環境・ニーズや条件の多様性がより広がっていくと想定されるため、多様性を考慮したより現実的かつ困難な設定に対応する技術**

- 多様・不均衡なデータから高精度でロバストな機械学習を可能にするマルチモーダルAI技術
- 少量のデータで基盤モデルを効果的に適応化させるzero-shot/few-shot学習・転移学習技術
- 今後増加が見込まれるロボットや車を含めた多数かつ多様なエッジ機器とクラウドの間で高効率な分散機械学習を可能とするエッジAI技術
- エッジ側の個々や環境の多様性に対して、サーバにデータを集約することなく対応可能とする連合学習技術
- AIの判断根拠をAI利用者や開発者の多様性に合わせて分かりやすく提示する説明可能AI

■ **様々なエッジとクラウドの間を循環しながら学習を繰り返すことに伴う基盤モデルの品質の保証や向上に資する技術**

- 基盤モデルがいつ、どこで、だれに、どのようなデータを使ってどのように学習されたのか（モデルの系統）を追跡するプロベナンス／リネージ技術
- 既存環境・タスクの知識を保持しつつ新規環境・タスクを検知し逐次適応する継続学習技術
- 基盤モデルの開発者・利用者を含めた循環進化全体のシステムとして、アプリケーション・目的・システム階層などの違いを吸収し持続可能性を高める品質保証技術

■ **循環進化におけるセキュリティやプライバシー保護技術**

- アカウント管理、認証・認可とセキュリティポリシー管理、 誤操作や悪意によるデータやモデルの漏洩や整合性の毀損の防止、不正アクセスからの保護など

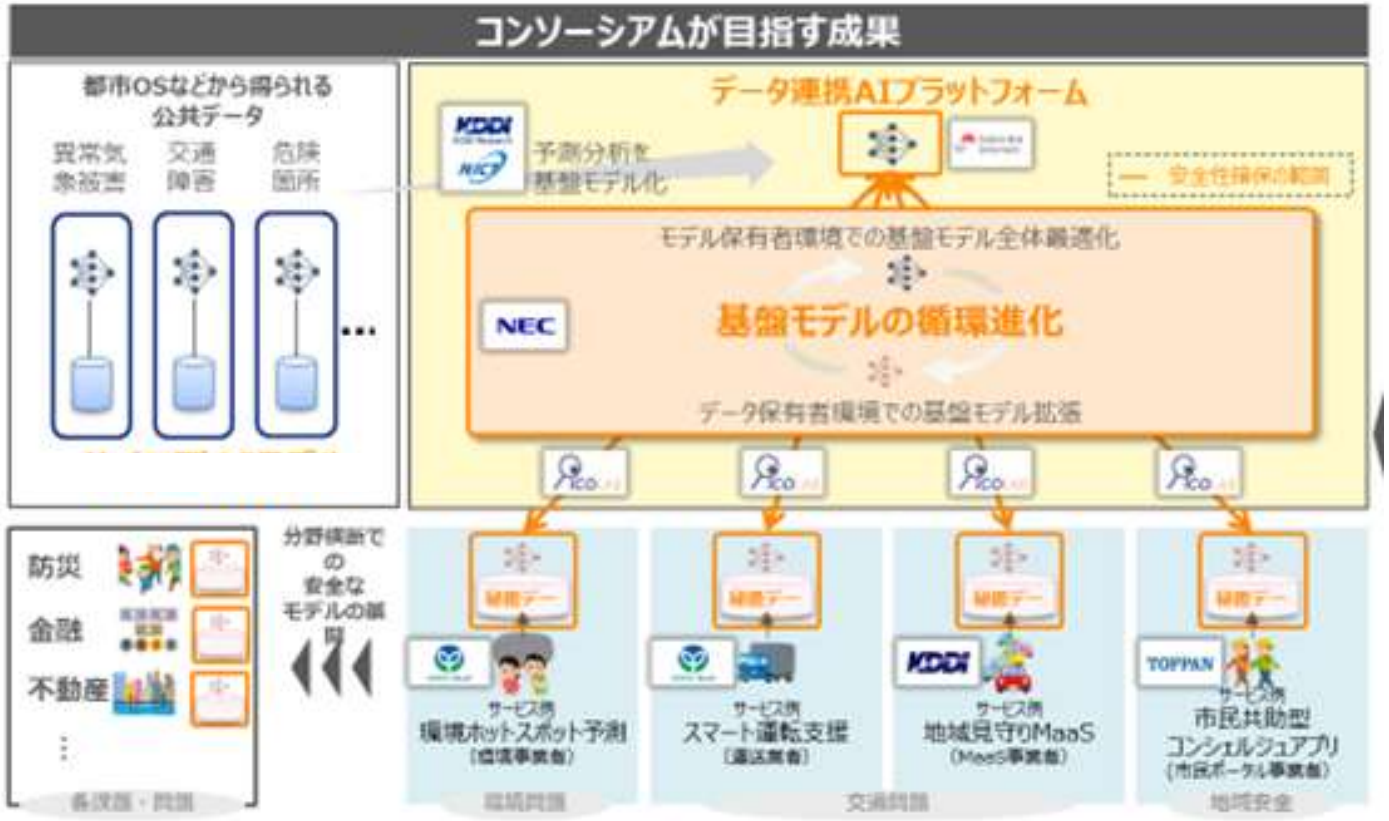
■ **アーキテクチャの標準化や参照システムの公開等を通じた循環進化に対する国際的なコンセンサス形成**

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

現在の取り組み

■ 総務省受託「安全なデータ連携による最適化 AI 」

- **循環進化の先行プラクティス**として、2023年7月より、**安心・安全・快適に移動できるまち**を、自治体・事業者・住民の共助により実現するための技術開発と社会実証に、10社のコンソーシアムで取り組み中



【マルチモーダルA】
- 多様・不均衡・少量データに対するロバストな深層学習技術

【エッジAI】
- エッジの多様性を考慮した高効率な連合型エッジAI技術、オフロード分散学習方式

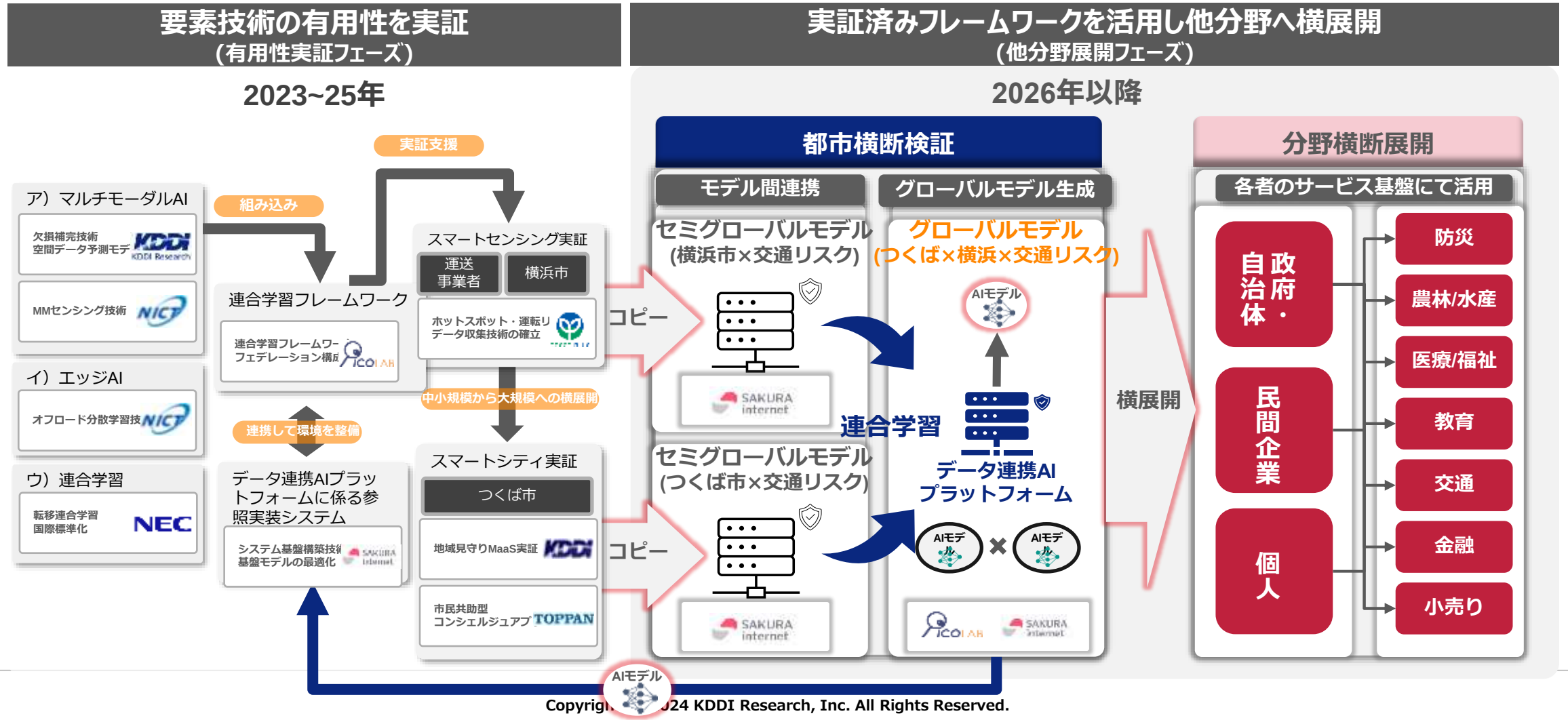
【連合学習】
- 安全に個別環境適応が可能なパーソナライズド連合学習技術、転移連合学習技術

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

現在の取り組み

■ 総務省受託「安全なデータ連携による最適化 AI 」

- 複数のタスク（交通リスク、街の安全リスクなど）に対し、それぞれ個別に循環進化の仕組みを構築
- プロジェクト終了の2025年までに要素技術の開発、複数の社会実証(スマートセンシング/シティ）を完了
- 2026年以降に、他都市への展開、分野横断展開を目指す



Copyright © 2024 KDDI Research, Inc. All Rights Reserved.

まとめ

- 現在の潮流と今後の見通しを受け、様々な組織・デバイスに散在するデータ・計算資源を安全・効率的に活用する**分散型AI**の取り組みが重要

- 今後の方向性として、日本の強みを活かした「AI基盤モデルの循環進化」を提示

- 実現のためには、データ・環境・ニーズの多様性に対応可能な分析技術、エッジ～クラウド間を循環しながら学習する際の品質向上に資する技術、循環進化におけるセキュリティ・プライバシー保護技術、国際的なコンセンサス形成　などが必要

- 循環進化の先行事例として、2023年7月より、安心・安全・快適に移動できるまちのための安全なデータ連携による最適化 AI の技術開発・社会実証に取り組み中

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.

Copyright(C) 2024 KDDI Research, Inc. All Rights Reserved.