総務省AIネットワーク社会推進会議
第5回議長ヒアリング

AIサプライチェーンに対する富士通の取組

2022年3月24日

富士通株式会社

ＡＩ倫理ガバナンス室　室長 荒堀 淳一
人工知能研究所　所長 穴井 宏和

AI倫理ガバナンス室の新設について（2022/2/1）


Sustainable Manufacturing


Digital Shifts


Consumer Experience


Business Applications


Healthy Living


Hybrid IT


Trusted Society

サステナブルな世界を実現する 7 Key Focus Areas

自社規律・社内外浸透

AI倫理ガバナンス室

最先端テクノロジーの研究・開発・実装にまつわる倫理に関する国際的な動向、政策、法制度の動向などを踏まえ、AI倫理ガバナンスに関する全社的かつ総合的な取り組みをさらに強化するため、社長直下に新設

研究開発

人工知能研究所・AI倫理研究センター

サステナブルな世界の実現のために、最先端のAI技術を国内外のアカデミアと連携して開発し、AI倫理のルール作りや課題への取組みを国内外団体と議論・主導し、ユーザ企業や社会の共感を喚起

AI領域における国内・国外連携

FUJITSU
富士通

関係府省庁

- 内閣府：AI戦略、AI人材育成
- 総務省：AIネットワーク社会
- 経産省：AI品質（産総研）

産業団体

- 経団連
- JEITA
- AIプロダクト品質保証コンソーシアム

研究機関

- 九州大学 ADS育成コンソーシアム
- 国立がん研究センター
- 理研API

国際組織

- AI4People
- OECD
- GPAI

標準化・産業団体

- ISO/IEC
- ECAI
- DigitalEurope
- JBCE
- ITI

研究機関

- 米 MIT、CMU
- 英 Oxford大学
- 独 ミュンヘン工科大学
- 仏 INRIA
- イスラエル ベングリオン大学

ほか

目 次（ エグゼクティブサマリ ）

01. AIサプライチェーンについて

AIサプライチェーンと既存サプライチェーンの関係

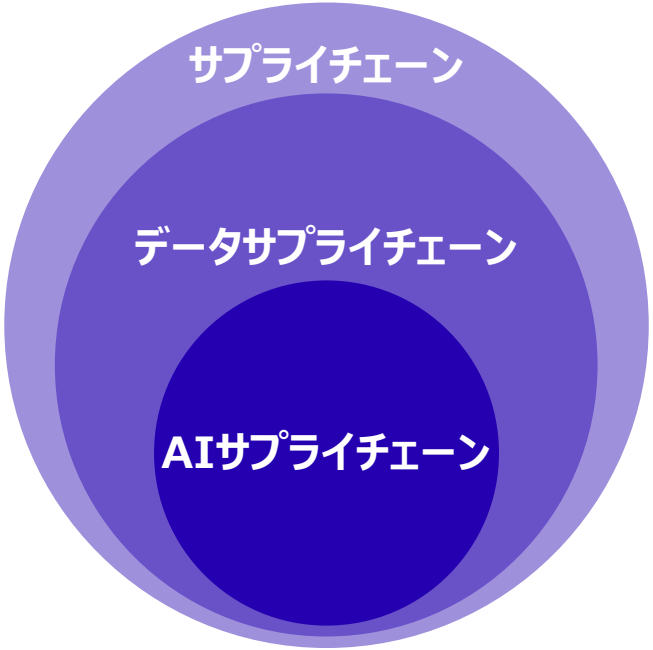
FUJITSU

AIサプライチェーンは、データサプライチェーンにAI実現のための要素を加えたもの

既存のサプライチェーン、データサプライチェーンに関わる法律やガイドライン、標準、技術を踏襲

AIサプライチェーン固有の要素

- **AIを含んだサプライチェーンのパターン**
- **AIサプライチェーンならではの技術**

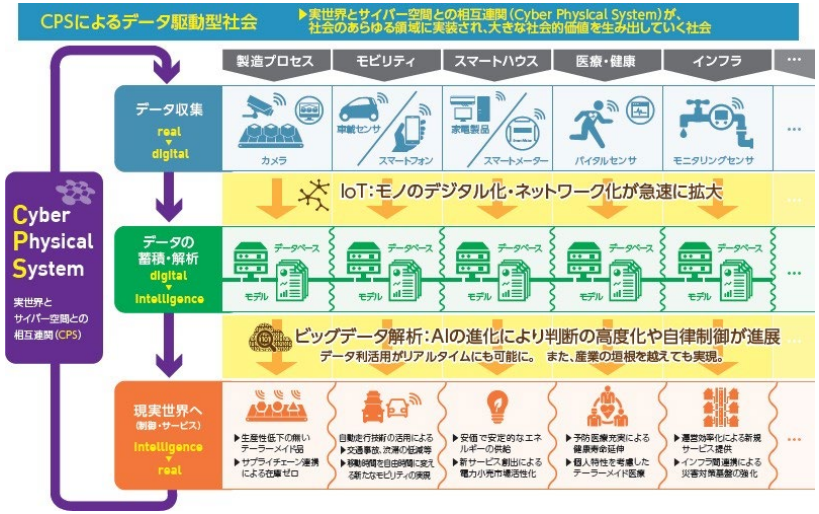


※ 現時点で、「AIサプライチェーン」という用語は確立されていない
”AI Supply Chain”で検索するとサプライチェーンをAIで効率化する話が主

【参考】データサプライチェーン

FUJITSU

データ収集 › データ蓄積・解析 › 制御・サービス



【図の出典】
経済産業省 産業構造審議会商務流通情報分科会 情報経済小委員会
「中間取りまとめ　～CPSによるデータ駆動型社会の到来を見据えた変革～」2015年5月21日

https://www.meti.go.jp/committee/sankoushin/shojo/johokeizai/report_001

【参考記事】
impress社 IT Leaders「データの信頼性確保に資するデータサプライチェーンとデータマネジメント：第5回」2021年10月15日

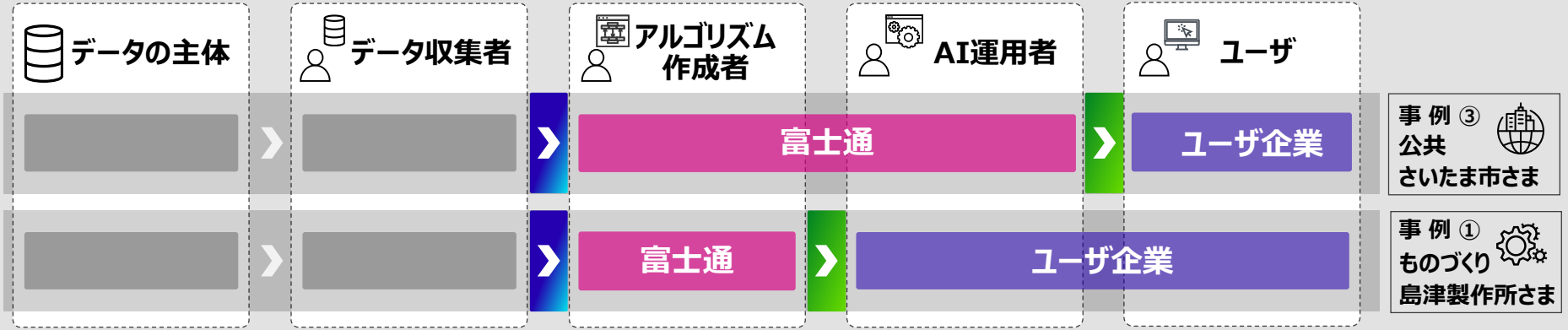
https://it.impress.co.jp/articles/-/22191

AIサプライチェーンの典型パターン (1/2)

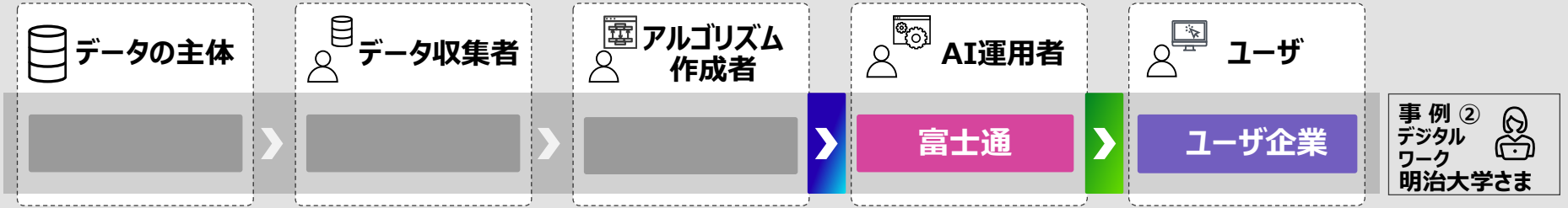
FUJITSU

パターン1（アルゴリズムの流通）

AIアルゴリズムを富士通が他社に提供するケース



AIアルゴリズムを富士通が他社から購入するケース

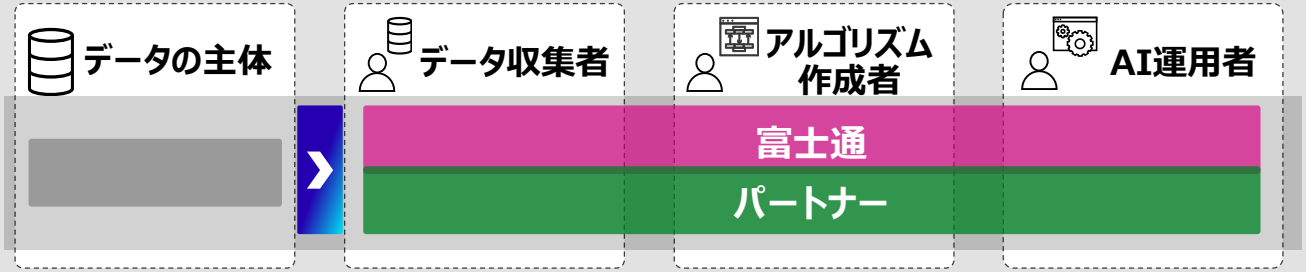


AIサプライチェーンの典型パターン (2/2)

FUJITSU

パターン2（データの流通）

共同研究



データを購入し、ユーザに転売するケース

AIサプライチェーンの法律と契約

FUJITSU

サプライチェーンにおける法律上の要請

1. ＣＳＲ調達（ 強制労働児童虐待の禁止、紛争鉱物の排除等 ）
2. グリーン調達 （ ISO14001 等 ）
3. 反社条項
4. 輸出管理 ： 米国輸出管理規則（ EAR ）等
5. 知的財産の保護
6. データ移転の制限（ 個人情報、ＧＤＰＲ 等 ）
7. 下請法
8. 独占禁止法 （ カルテル・談合 等 ）

サプライチェーンにおける契約上の処置例

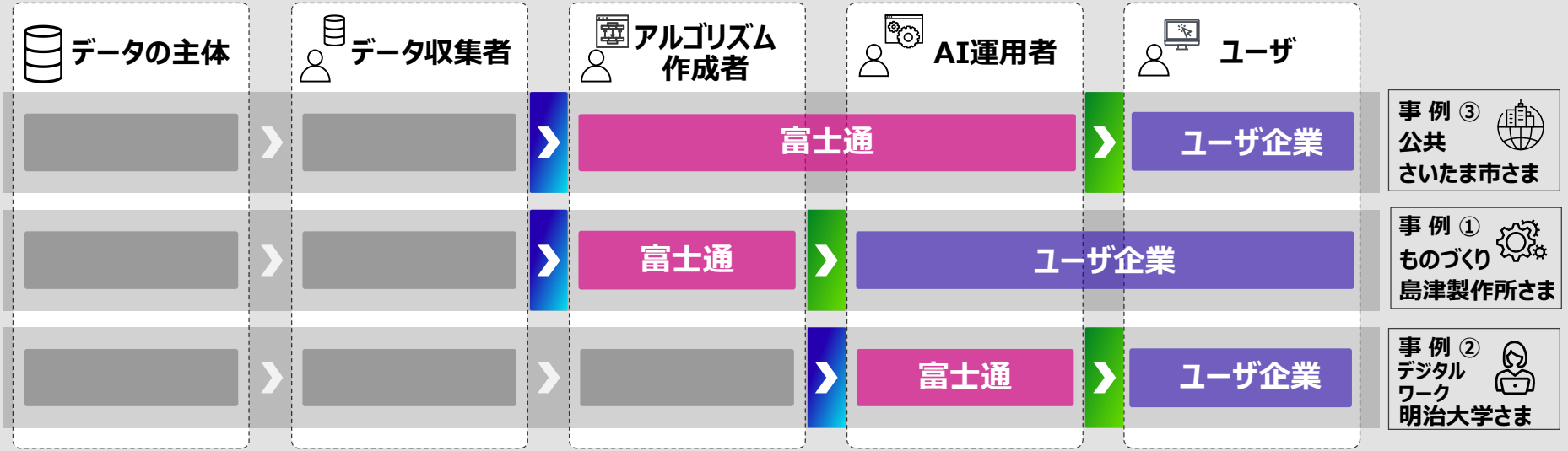
1. **法律上の要請**
   - 前頁の法律上の要請を遵守
2. **契約による担保**
   - ライセンス使用制限
   - 報告義務
   - 監査受入れ義務
   - 表明保証
   - 損害賠償による救済

AIサプライチェーンにおける特有の法的課題は現状では認識しない

典型パターンに対する契約上の措置 (1/2)

FUJITSU

パターン1（アルゴリズムの流通）



当社は、データ収集者の正当な権限等を契約で表明保証させ、必要により監査をおこなう

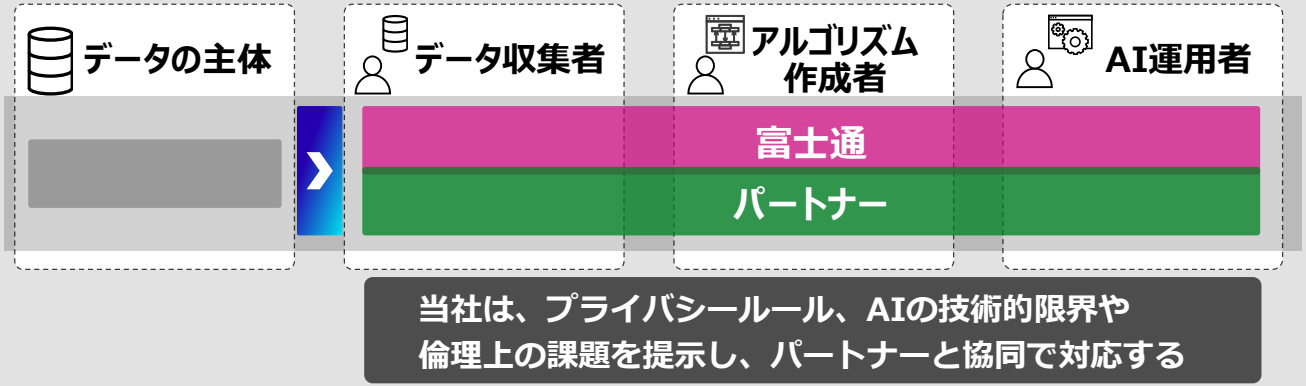
当社は、AIの技術的限界や倫理上の課題を提示し、ユーザーと協同で対応する

典型パターンに対する契約上の措置 (2/2)

FUJITSU

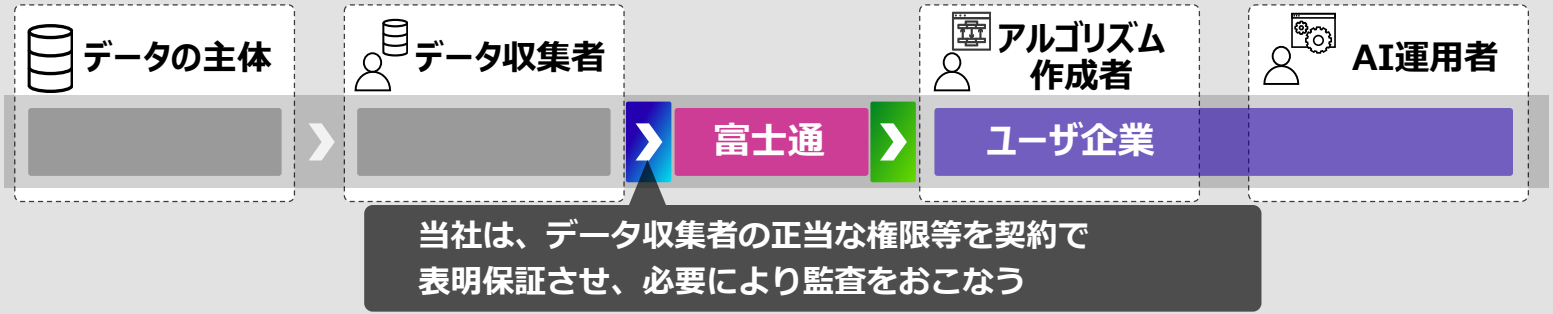
パターン2（データの流通）

共同研究



なお、契約に加え、技術によるデータ信用性の担保も考えられる

データを購入し、ユーザに転売するケース



02. 富士通のAIビジネス事例にみるAIサプライチェーンの実例

事例① 質量分析計の測定データを高精度に自動解析

FUJITSU

活用シーン ものづくり

島津製作所 さま

熟練者でなくてもばらつきのない高精度な分析作業が短時間で可能に

サービス: Zinrai活用支援サービス | 技術: ディープラーニング | 商品販売: 2019年9月

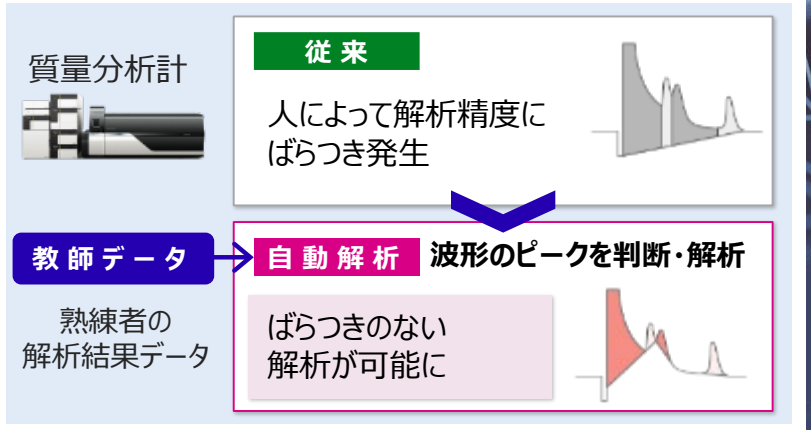
課題・技術・効果

課題
計測機器から得られる計測結果の解析に時間がかかる
作業者によって解析精度にばらつきが発生する

技術
熟練者が行った解析結果をAIに学習させることで、熟練者と同等レベルの解析を実現
従来のアルゴリズムでは作業者が行っていたパラメータ調整をAIが自動で最適化

効果
従来のアルゴリズムと比較して作業にかかる時間を約1/3に短縮し熟練者でなくてもばらつきのない解析が可能に

活用シーン



サプライチェーンパターン：AI部品提供

FUJITSU

データの主体	データ収集者	アルゴリズム作成者	AI運用者	ユーザー
島津製作所さま		富士通	島津製作所さま	

事例①
ものづくり
島津製作所さま

島津製作所さま

サンプル

質量分析計

クロマトグラムデータ

データの蓄積・維持・管理

データへのアノテーション

富士通

共同開発で共有

AIモデル開発

共同でモデルをブラッシュアップ

学習アルゴリズム

AIモデル

AIソフトウェア開発

ライセンス提供

データ解析製品

全世界へ販売

データ解析業務で利用

世界各国のエンドユーザさま

要 点

FUJITSU

島津製作所さまとの共同研究から展開した協業プロジェクト

- 2016年から共同研究開始
- 研究成果を2019年に製品化

データの観点

- 本製品開発のために、島津製作所さまがデータを収集
  AI開発用のデータへのアノテーションまでを実施
- 共同研究、共同開発の契約の元、富士通にデータを共有

システム開発の観点

- 富士通は島津製作所さまデータを用い、AIモデルを開発
  [学習データは、クロマトグラムデータのみ（物質名などの情報は含まれず）]
  [島津製作所さまと打ち合わせをしながらAIモデルをブラッシュアップ]
- 当該モデルを組み込んだAIソフトウェアを富士通が開発、島津製作所さまに提供
- 島津製作所さまが、富士通のAIを組み込みデータ解析ソフトウェアを開発

AIソフトウェア・AIモデルのエンドユーザへの提供

- データ解析ソフトウェアを島津製作所さまが全世界に販売
- カスタム対応が必要な顧客に対しては、別途カスタムAIモデルを構築し、
  製品の個別オプションとして提供

事例② 職員約1,000人がAI翻訳サービスの利用を開始

FUJITSU

活用シーン
デジタルワーク

明治大学 さま

働き方改革と翻訳品質向上への挑戦

製品: Zinrai Translation Service

技術: 文書翻訳

課題・技術・効果

課 題

外国人留学生や海外協定校の増加に伴い、
各種掲示物や協定書等の翻訳業務が急増
作業負荷の高まりや翻訳品質が課題

技 術

ニューラル機械翻訳エンジン*による高精度な翻訳
（日本語⇔英語、日本語⇔中国語）
ファイルをそのまま翻訳(PowerPoint、Word、Excel、PDF)
逆翻訳機能による翻訳精度確認の作業を効率化

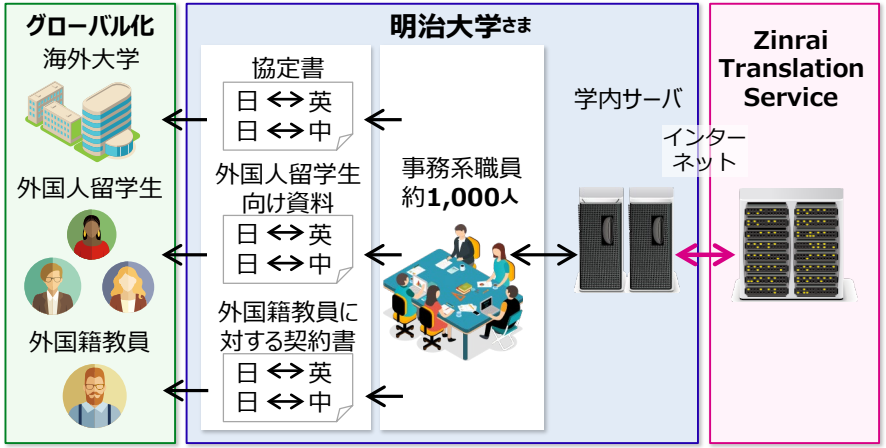
効 果

翻訳作業の負荷軽減と翻訳品質の向上
約800語の明治大学固有の名詞や専門用語を
辞書に登録し、表現を統一化

(*)株式会社みらい翻訳が国立研究開発法人情報通信研究機構（NICT）と共同開発した、
ニューラル機械翻訳（NMT）技術とみらい翻訳プラットフォーム技術を利用しています

活用シーン

翻訳作業にかかっていた時間を大幅に短縮



定額制の料金プラン* 画面カスタマイズ逆翻訳の機能追加 (*)16.5万円/月(税別)～
3つの料金プラン

サプライチェーンパターン：AIサービス提供

FUJITSU

データの主体	データ収集者	アルゴリズム作成者	AI運用者	ユーザー
みらい翻訳さま			富士通	明治大学さま

事例②
デジタルワーク
明治大学さま

みらい翻訳さま

AIモデルの維持・管理　　データの蓄積・維持・管理

学習アルゴリズム

AI翻訳エンジン開発*

*株式会社みらい翻訳が国立研究開発法人情報通信研究機構（NICT）と共同開発した、ニューラル機械翻訳（NMT）技術とみらい翻訳プラットフォーム技術を利用しています。

明治大学さま

文書翻訳業務で利用

富士通

ライセンシング

クラウド翻訳サービス開発

クラウドサービス提供

富士通の情報管理規定に従い顧客からのデータは厳重に管理

翻訳後オリジナル文書は破棄

要 点

FUJITSU

みらい翻訳さまとの協業プロジェクト

- みらい翻訳さまの製品を富士通のサービスに組み込み、富士通が販売

データの観点

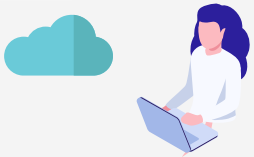
- 製品開発段階でみらい翻訳さまが学習用データを利用
  [富士通は学習用データにはアクセスできない]
- エンドユーザさまが翻訳したいファイルを富士通側クラウドサービスにアップロード
  翻訳したデータをエンドユーザさまがダウンロード
  翻訳が終わり次第、エンドユーザさまからアップロードされたデータ、翻訳結果は削除

システム開発の観点

- 富士通はみらい翻訳さまのサービスを組み込んだ形で自社サービスを提供。
- 翻訳の品質については、みらい翻訳さまが担保

AIソフトウェアのエンドユーザへの提供

- エンドユーザさまはクラウドサービスとして利用するのみ
  AIソフトウェア、AIモデルに直接触れることはない



事例③ 保育所入所選考の自動最適化

FUJITSU

活用シーン 公共

さいたま市 さま

非常に多くの日数のかかる複雑な保育所の入所割り当ても わずか数秒で最適解を自動算出

製品:MICJET MISALIO 子ども・子育て支援保育所AI入所選考

課題・技術・効果

課 題

「きょうだいで同じ保育所を希望」「別々の保育所でも良いが、きょうだいの片方しか入れないのなら入所しない」などの多様な希望を満たす割り当てに時間がかかる

技 術

ゲーム理論を元にした独自のマッチング技術により、申請者の希望を**最大限満足する割り当て**を実現

効 果

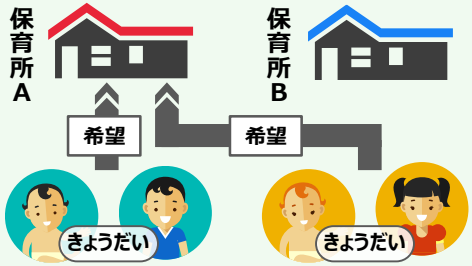
20-30名の職員が非常に多くの日数をかけて実施していた割り当て作業をわずか数秒で実現

自治体向け保育業務支援システム「MICJET MISALIO 子ども・子育て支援保育所AI入所選考」として2018年11月より提供

活用シーン

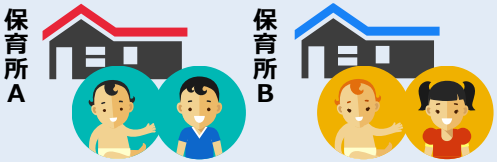
希 望

2組のどちらのきょうだいも**できるだけ同じ保育所**で保育所Aを好む・・・



最 適 解

「優先順位の高い人」の**「出来る限り高い希望」**をかなえる



サプライチェーンパターン：AIアルゴリズム提供

FUJITSU

データの主体	データ収集者	アルゴリズム作成者	AI運用者	ユーザー
さいたま市さま		富士通		他自治体

事例③ 公共 さいたま市さま

さいたま市さま データの蓄積・維持・管理

保育所関連システム　入所希望者データ

難易度の高い要求仕様

さいたま市職員さま

実証実験で共有

入所マッチング結果検証

人工データ

富士通

AIアルゴリズム開発

AIアルゴリズム

富士通の情報管理規定に従い厳重に管理

富士通研究部門

AIアルゴリズム供給

製品開発

他自治体

保育所AI入所選考ソフト

販売

保育所AI入所選考ソフト

富士通事業化部門

児童への保育所割り当て

住民

要 点

FUJITSU

富士通の研究部門とさいたま市さまとの実証プロジェクトから製品化に至った事例

- 富士通と九州大学さまの数学の研究者が、児童を保育所に最適に割り振るアルゴリズムを開発
- さいたま市さまとの実証実験で当該アルゴリズムの有効性を確認
- 実証実験の結果を受けて、当該アルゴリズムを搭載した保育所AI入所選考ソフトを富士通が販売

データの観点

- 本アルゴリズムの有効性検証のために、さいたま市さまから実証実験の目的のためにデータを共有
- 統計的機械学習型のAIアルゴリズムでなく、数理最適化型であるため、製品に住民データから作られたモデルが搭載されることはない

システム開発の観点

- 実証実験で有効性を確認できたアルゴリズムを研究部門から事業化部門に移管
- 事業化部門は当該アルゴリズムを製品に組み込み。
  自治体の難易度の高い要求仕様（公平性の担保、プライバシー保護など）に応えるロジックを開発、提供

AIソフトウェア・AIモデルのエンドユーザへの提供

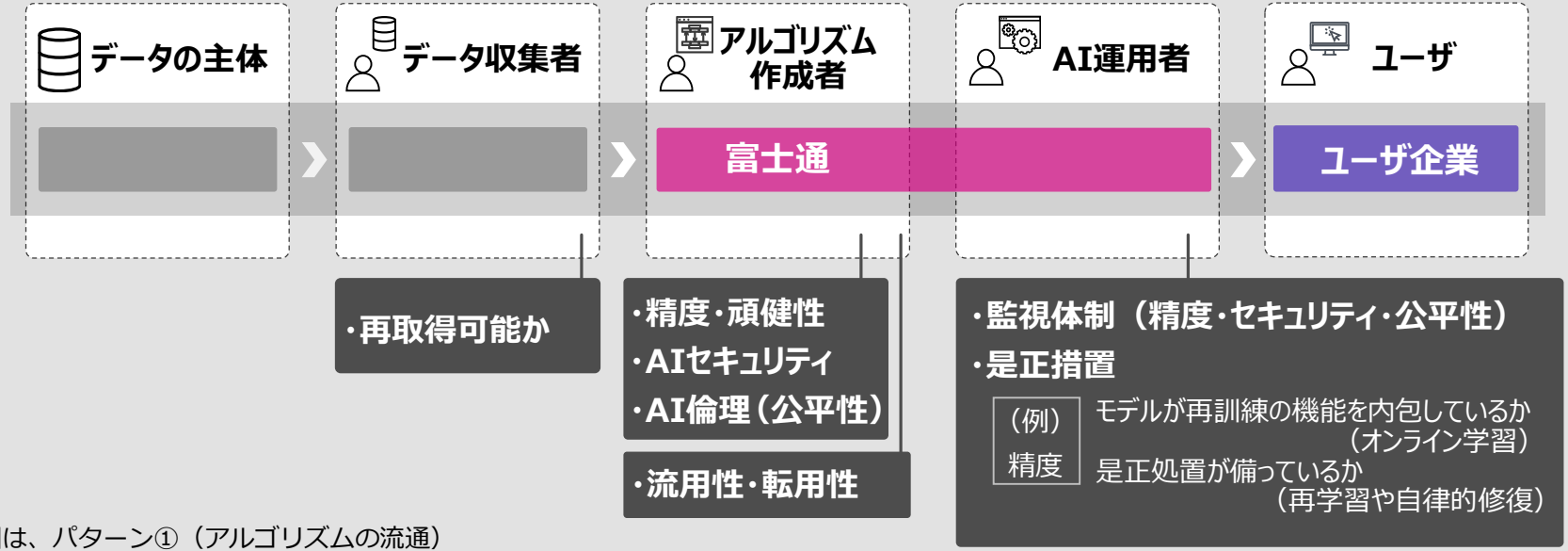
- AI機能が搭載された保育所AI入所選考ソフトを富士通が自治体に販売
- 自治体職員が、当該ソフトウェアを用いて、入所希望者の多様な希望
  （例：きょうだいが同じ保育園が良い、上の子だけでも優先して入所させたい、など）を基に、入所選考結果を導出

FUJITSU

03. 今後の論点
（AIサプライチェーンを支える技術）

将来像：AIサプライチェーン固有の要素

- 各ステークホルダー間の境界に関わる要素について、AI固有の監査が必要
- 運用中の変化が課題となるため、「AI運用者」の役割が重要
- ライフサイクル（反復プロセス）の扱いが肝要　例えば、データの再取得が可能な契約になっているか、等



※ 図は、パターン①（アルゴリズムの流通）
AIアルゴリズムを富士通が他社に提供するケースの場合

AI運用時の品質保証における課題（1）

FUJITSU

開発時の学習データから運用時の入力データの傾向や外部環境が変化することで、
AIモデルが陳腐化（運用時のモデル精度低下）

精度低下を放置すると大きな損害を招く危険性が高まり、
それを防ぐには、定期的な正解付けにより、精度の監視が必要

AI運用時の品質保証における課題（2）

FUJITSU

課題 1	精度低下や異常状態を見逃すリスクがある	**入力データの変化傾向を追跡して、AIの精度を自動推定**	精度監視
課題 2	高精度を維持するために頻繁なメンテナンスが必要	**既存モデルの再学習なく、精度低下を抑制**	自動修復
課題 3	再学習に膨大なコストがかかる	**自動ラベル推定により、正解付けが必要なデータ数を削減**	再学習支援

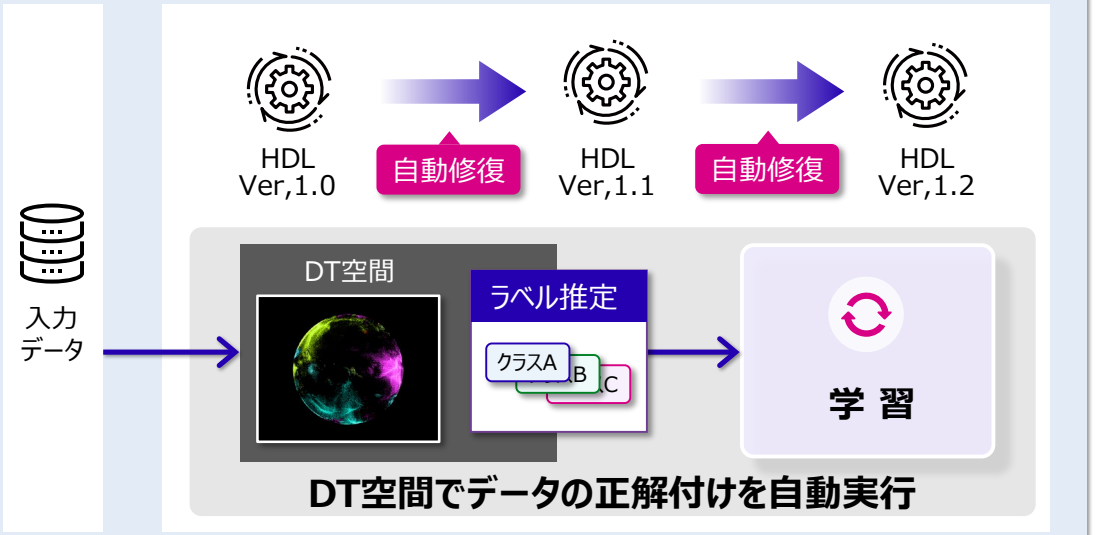
AI品質技術: HDL（High Durability Learning）

FUJITSU

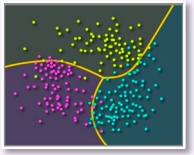
HDL = 高耐性学習

AI運用時の正解付け・精度低下の監視・モデル修復を自動化

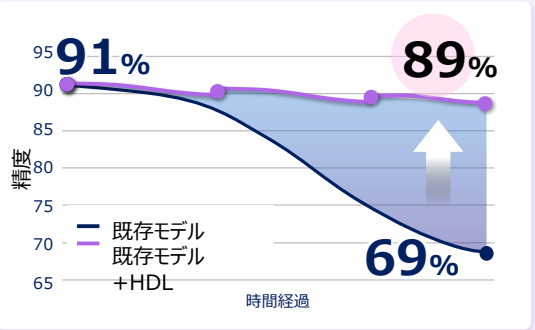
HDLモデルの自動修復



精度低下の自動修復



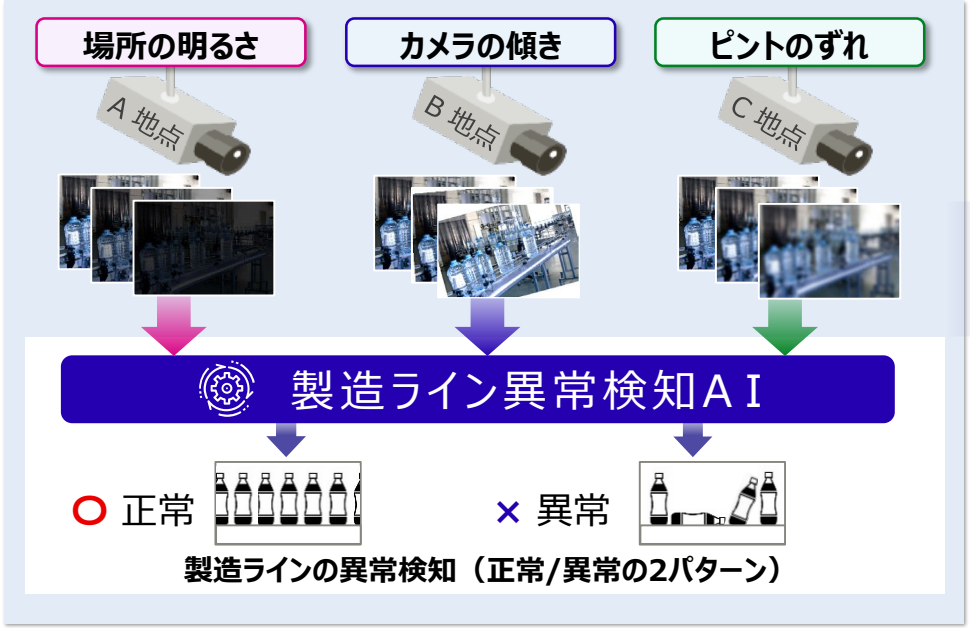
精度低下・
異常状態による
リスクを軽減



HDLの適用事例（製造ラインでの異常検知AI）

FUJITSU

1. 飲料品生産工場の製造ラインにおける画像検査AIに適用
2. 現場で想定される**多様な環境変化**に対して、**高耐性なAI運用を実現**



High Durability Learning 適用

製造ライン異常検知

① 検知精度（AUC）	**0.21**ポイント向上
② 自動修復（Recall）	**70**% → **98**%
③ 正解付け工数	**96**% 削減

まとめ（再掲）

01. AIサプライチェーンについて

既存のデータサプライチェーンの枠組みの延長線上に成り立つ

既存の法律面・契約面を踏襲して運用可能

AIならではの追加要素は技術面で対応するのが適当

02. 富士通のAIビジネス事例にみるAIサプライチェーンの実例

典型例。＋お客様がAIアルゴリズムを再販する例　【 島津様事例（計測機器データの自動解析）】

AIアルゴリズムを他社様から調達する例　【 明治大学様事例（文書翻訳）】

公共からの厳しい要求(公平性、データ漏洩防止、プライバシー保護)の例　【 さいたま市様事例（保育所入所選考）】

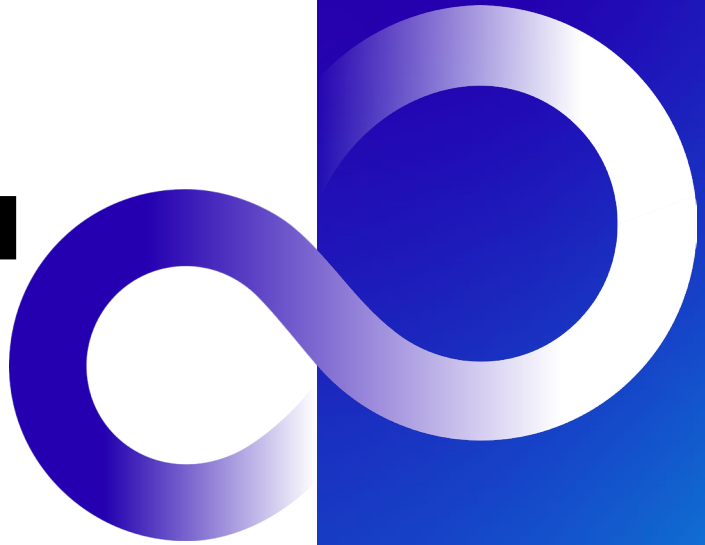
03. 今後の論点（AIサプライチェーンを支える技術）

AIの品質保証、AIセキュリティ、AI倫理（公平性）、運用時監視

AIの運用時の精度低下を監視し、修復する技術HDL

データとAIモデルへの攻撃に対するセキュリティ技術

Thank you



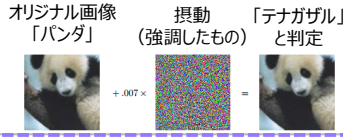
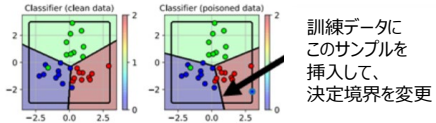
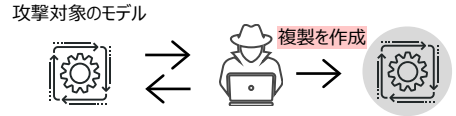
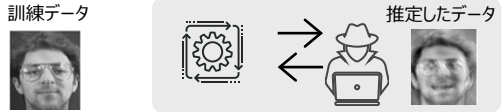
FUJITSU

富士通研究所におけるAIセキュリティの取り組み

FUJITSU

AI特有のサプライチェーンセキュリティ対策に関連する研究

・訓練データ・モデルがAI特有の要素
・強くサプライチェーンに関連する２攻撃の対策技術を紹介（次頁以降）

脅威	分類	攻撃イメージ
だます	**敵対的サンプル** **Adversarial Example**	
	訓練データ・モデル汚染 **Poisoning**	
情報を盗む	**モデル抽出** **Model Extraction**	
	訓練データ復元 **Model Inversion**	

想定される脅威例

データ提供元・開発元における**攻撃埋め込み**

運用委託先からの**情報流出**

汚染攻撃対策技術

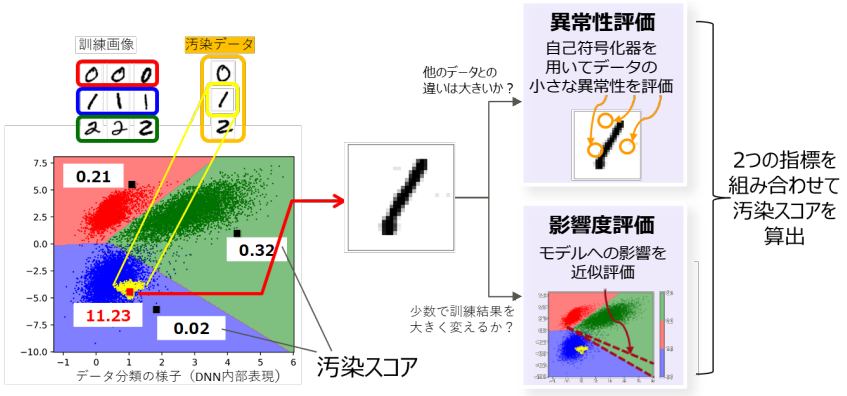
訓練データ･モデルの汚染攻撃（Poisoning）への対策

応用例：提供された訓練データの検査、納品されたモデル･OSSモデルの検査

訓練データ中の汚染データを検知

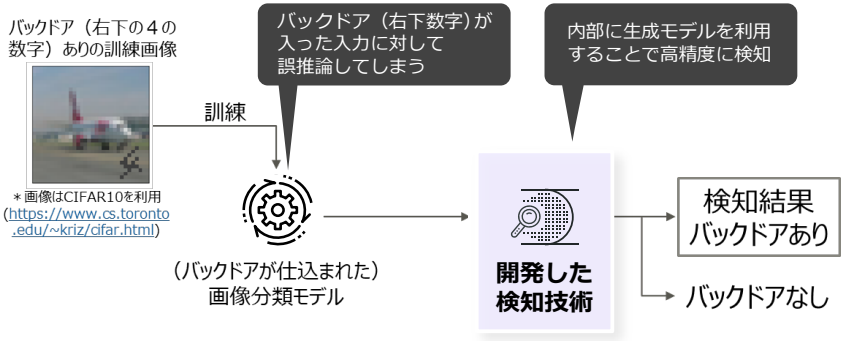
訓練データセットのすべてのデータに対して
汚染度を表すスコアを算出

文字認識Deep Neural Network(DNN)訓練データの汚染検知例



モデルに仕込まれたバックドアを検出

モデル内のバックドア（特定の印を含む画像データを
誤推論する仕組み）を発見



訓練データ復元攻撃対策技術

FUJITSU

訓練データ復元攻撃（Model Inversion）への対策

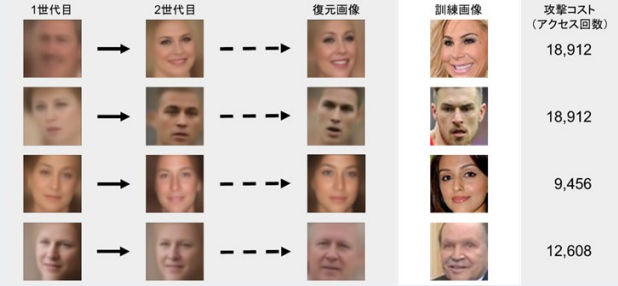
応用例：システム運用先でモデルから訓練データ窃取されないための対策

モデルへの訓練データ復元攻撃を妨害

訓練データに特殊なデータを追加することで、モデル性能の低下をおさえつつ復元攻撃耐性を向上

攻撃耐性の評価

復元攻撃性を評価
例：顔認識モデル



防御技術の適用

復元攻撃性を向上
例：手書き文字認識モデル

	通常のモデル	提案技術で訓練したモデル	従来技術（DP）
モデル性能（Accuracy）	0.9888	0.9863	0.7420
訓練データ推定に対する耐性（攻撃結果）	復元成功 「5」: 「6」: 「8」:	復元失敗 「5」: 「6」: 「8」:	復元失敗 「5」: 「6」: 「8」:

性能低下を改善

約**24.7**%

↓

約**0.25**%