

理研AIPセンター： これまでの研究成果と 今後の方向性について

センター長 杉山 将



<https://aip.riken.jp/>



AIPセンタービジョン

機械学習
の応用

2000年代
IT企業

連携

Society 5.0
ビッグデータ
ビジネス&サイエンス

AIP

未来の
AI社会

機械学習
の技術

1990年代
統計的学習
(SVM, Bayes)

目的

2010年代
神経科学的学習
(深層学習)

汎用

近未来
限定情報
機械学習

AIP



Cyber-space
機械学習

Cyber-physical
機械学習

機械学習
everywhere

- AIP常設研究グループ
 - 目的指向基盤技術研究グループ(上田)
 - 社会における人工知能研究グループ(橋田)
 - 汎用基盤技術研究グループ(杉山)



- AIP連携センター
 - NEC
 - 東芝
 - 富士通

ご説明の流れ



1. これまでの研究成果:

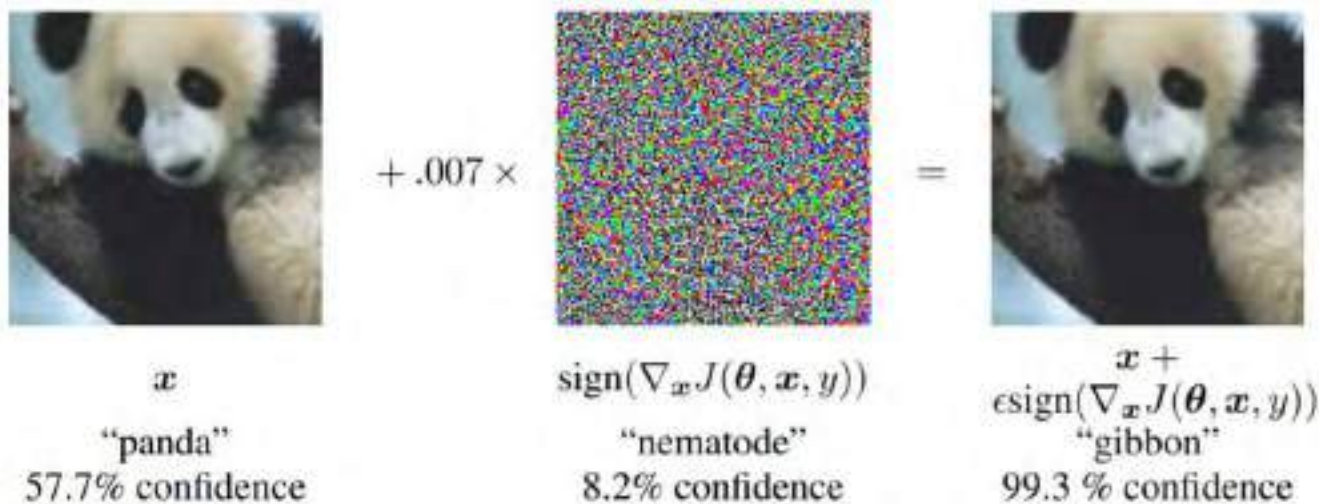
- A) 汎用基盤技術研究グループ(汎用G)
- B) 目的指向基盤技術研究グループ(目的G)
- C) 社会における人工知能研究グループ(社会G)

2. 今後の方向性

■ 背景: 現在主流の深層学習には様々な弱点がある

- 信頼性に対する脆弱性, 高データ収集コスト, 膨大な学習時間, 極度の忘却性...

Goodfellow et al.
(ICLR2015)



■ 目標: 日本発の独自のAI基盤技術を創出

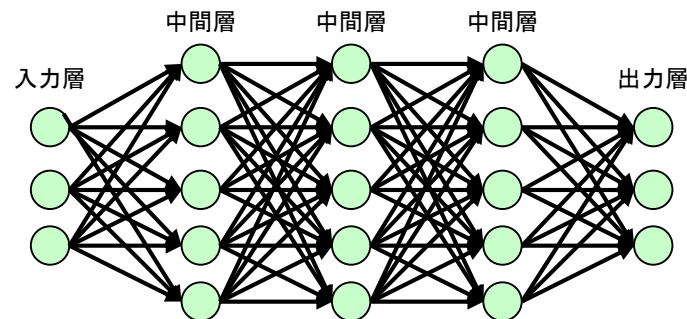
- 深層学習の原理を**数学的に解明**し, 更なる性能向上へ
- 上記難題を解決する**次世代の機械学習手法**の開発

成果1-1: 深層学習の理論

5

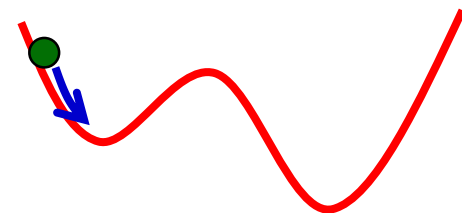
■ 深層学習:

- 多数の中間層を積み重ねる
- 最適化は非常に難しい
- 高次元データに対しても予測性能が高い



■ 研究成果: 深層学習が本当に優れていることを数学的に証明

- 大域的最適化が可能
Suzuki & Akiyama (ICLR2021)
- 高次元でも予測が当たる
Suzuki (NeurIPS2020)
- 万能な近似能力がある
Teshima et al. (NeurIPS2020, oral)



$$\widehat{L}(X_k) - \int \widehat{L}(x) d\pi_\infty(x) \lesssim \exp(-\Lambda_\eta^* k \eta) + \frac{c_\beta}{\Lambda_0^*} \eta^{1/2-\kappa}$$

$$dX_t = -\nabla \left(\widehat{L}(X_t) + \frac{\lambda}{2} \|X_t\|_{\mathcal{H}_K}^2 \right) dt + \sqrt{\frac{2}{\beta}} d\xi_t$$

$$R_{\text{lin}}(\mathcal{F}_\gamma) \gtrsim n^{-\frac{2\tilde{\beta}+d}{2\tilde{\beta}+2d}-\kappa'}$$

$$R_{\text{lin}}(\mathcal{F}_\gamma) := \inf_{\widehat{f}: \text{linear}} \sup_{f^\circ \in \mathcal{F}_\gamma} \mathbb{E}_{D_n} [\|\widehat{f} - f^\circ\|_{L_2(P_X)}^2]$$

$$\mathbb{E}_{D^n} \left[\mathbb{E}_{W_k} [\|f_{W_k} - f^\circ\|_{L_2(P_X)}^2 | D_n] \right] \lesssim n^{-\frac{\gamma}{\alpha_1 - 3\alpha_2 + 1}} + \Xi_k$$

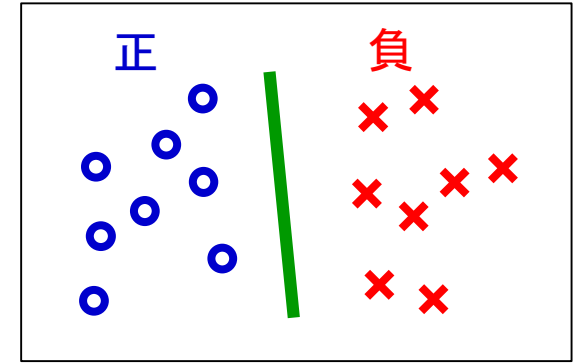
成果1-2: 弱教師データからの機械学習 6

■ 従来の深層学習は、「正しい」教師情報がついたデータが必要

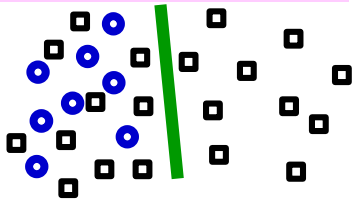
- 音声認識, 画像理解, 言語翻訳...

■ 簡単に集められる「弱い」教師データのみから学習できる理論体系を構築

- 医療, 生体, 生命, モビリティなどに対して実証実験中



正ラベルなし分類



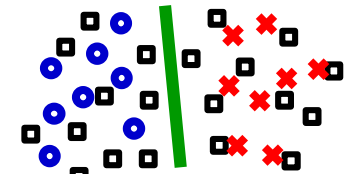
du Plessis, Niu & Sugiyama (NIPS2014, ICML2015)
Niu, du Plessis, Sakai, Ma & Sugiyama (NIPS2016)
Kiryo, du Plessis, Niu & Sugiyama (NIPS2017)
Hsieh, Niu & Sugiyama (ICML2019)

ラベルなしラベルなし分類



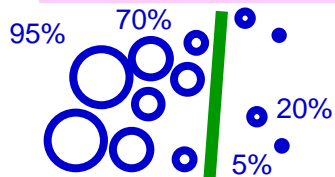
du Plessis, Niu & Sugiyama (TAAI2013)
Lu, Niu, Menon & Sugiyama (ICLR2019)
Charoenphakdee, Lee & Sugiyama (ICML2019)
Lu, Zhang, Niu & Sugiyama (AISTATS2020)

正負ラベルなし分類



Sakai, du Plessis, Niu & Sugiyama (ICML2017)
Sakai, Niu & Sugiyama (MLJ2018)

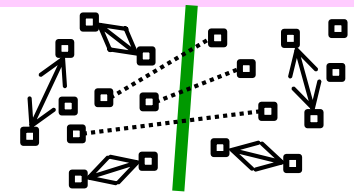
正信頼度学習



Ishida, Niu & Sugiyama (NeurIPS2018)

MIT Pressより弱教師付き学習の全貌を
まとめた専門書を出版(2021年出版予定)
Sugiyama, Bao, Ishida, Lu, Sakai & Niu,
Machine Learning from Weak Supervision

類似非類似ラベルなし分類



Bao, Niu & Sugiyama (ICML2018)
Shimada, Bao, Sato & Sugiyama (arXiv2019)
Dan, Bao & Sugiyama (arXiv2020)

成果1-3: 因果推論

■ 相関と因果は異なる:

- チョコ消費量からノーベル賞受賞者数が予測できる
- チョコ消費量を増やしてもノーベル賞は増えない

■ 因果推論: 介入せずに因果関係を知りたい

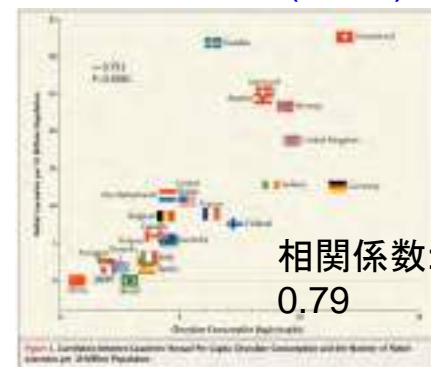
- 隠れた要因の扱いが,
因果推論の最大の課題の一つ

■ 隠れ要因の存在下でも, 全体構造が推定可能な初めての手法を開発:

- 音声の分離に用いられる技術を応用して, 隠れ要因を分離

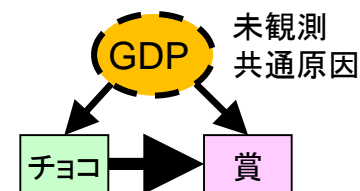
Maeda & Shimizu (AISTATS2020)

Messerli (2012)



ノーベル賞受賞者数

チョコレート消費量



ご説明の流れ



1. これまでの研究成果:

- A) 汎用基盤技術研究グループ(汎用G)
- B) 目的指向基盤技術研究グループ(目的G)
- C) 社会における人工知能研究グループ(社会G)

2. 今後の方向性

目的G

- **背景**: 日本は様々な社会課題に直面しているが、その解決に向けた科学研究の国際競争力は低下



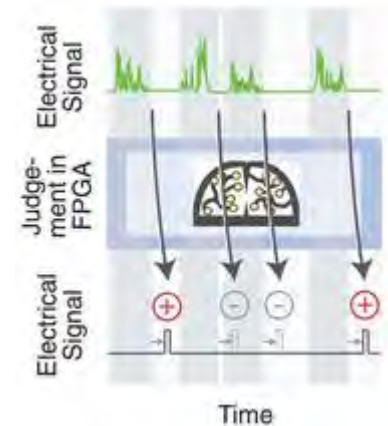
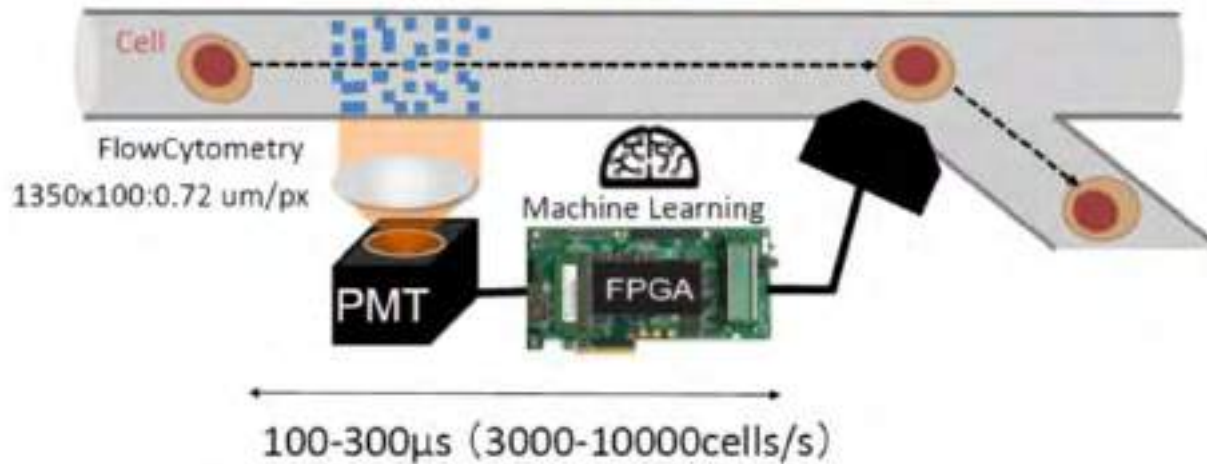
- **目標**: AIの社会実装・問題解決に貢献
 - 日本が強い科学研究分野を更に強化
(がん, 再生医療, 材料...)
 - 日本が直面している社会課題解決へ貢献
(防災減災, 高齢者ヘルスケア, 記述文自動評価...)

成果2-1: 細胞の自動仕分け

10

「ゴーストサイトメトリー」

Ota et al. (Science 2018)



- 細胞を流し, 正常なものと異常なものを自動選別:
 - 深層学習による画像認識では計算が遅すぎる
- 画像化せず**ゴースト情報**から**実時間分類**を実現:
 - シンクサイト(AIPメンバが創業)と日立が商品化に着手
 - 国立がん研究センターと腫瘍細胞分取への応用を開始

成果2-2: 前立腺病理画像からのがん検出 11

Yamamoto et al. (Nature Communications 2019)

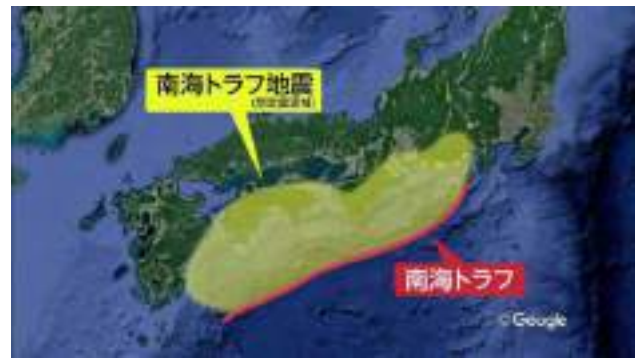
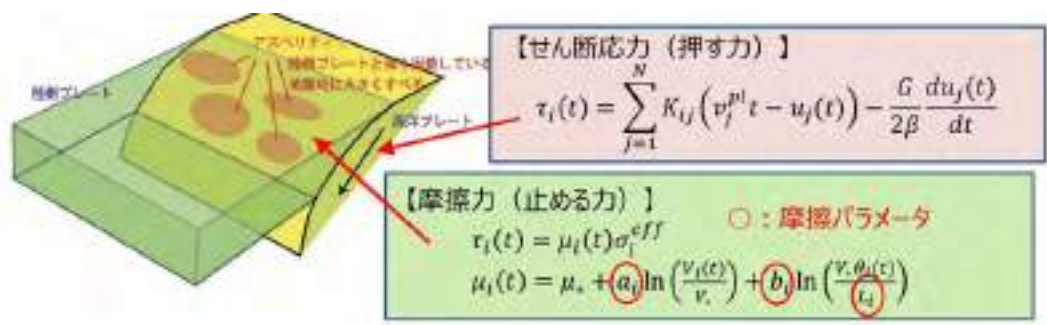
- 前立腺がんは男性のがんの最大の要因
- 教師付き分類では医師のラベル付けが困難
- 110億枚の病理画像パッチから、教師なし深層学習とクラスタリングにより特徴抽出
 - 従来のグリソンスコアに加えて、**がん領域以外の間質の変化**など、新たな特徴を発見
- さらなる発展：
 - 京大とiPS細胞に展開
 - 東大・北里大と白血病，乳癌に展開
 - 企業との実用化

癌再発予測(1年以内)	AUC
病理医 (グリソンスコアを使用) 	0.744
AI (AIが見出した新規特徴を使用) 	0.82
AIと病理医 (新規特徴+グリソンスコアを併用) 	0.842

成果2-3: 南海トラフ地震の周期予測 12

■これまでの研究により数理モデルが確立:

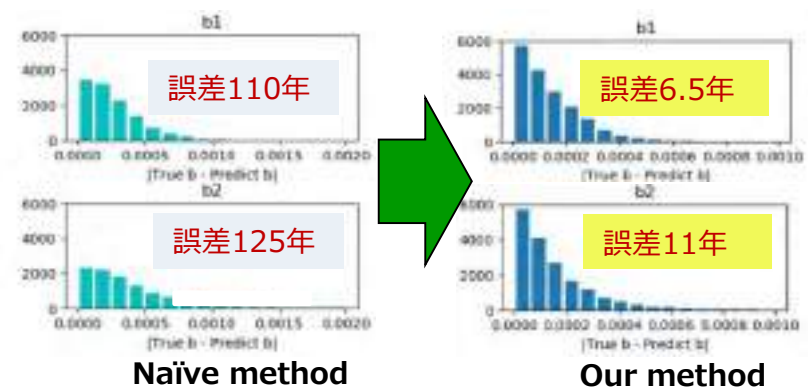
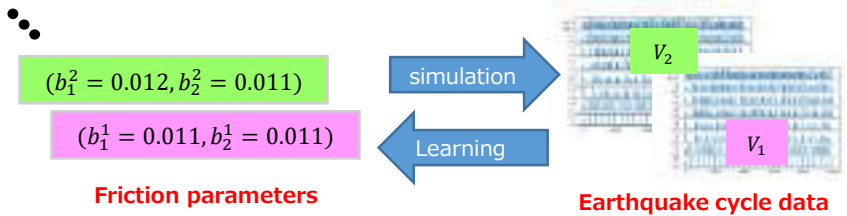
- しかし, パラメータの決定が困難
- 教師データも集めにくい



<https://www.fnn.jp/articles/-/22389>

■シミュレーションに基づく機械学習: Hachiya et al. (EGU2019)

- 疑似データ生成とパラメータ推定を繰り返す



成果2-4: 文章の自動評価

(Mizumoto et al., BEA2019; Funayama et al., ACL-SRW2020)

■ 文章の質を自動評価したい:

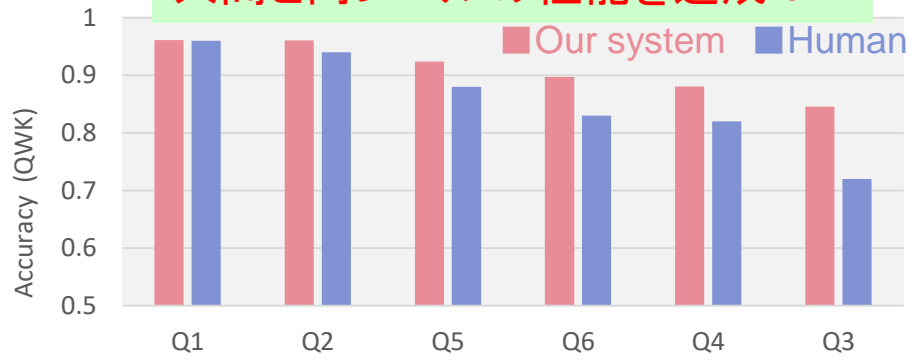
- 作文能力の向上, 作文評価の自動化

■ しかし, 大量に教師データを準備することは困難:

- 半教師付き学習, 自己学習などを活用
- 予測信頼度と事例提示により**説明性**を担保

インタラクティブな評価システム

200-400 の教師データだけから, 人間と同レベルの性能を達成!



日本語記述式問題自動採点デモ

次の文章を読んで、後の問いに答えよ。

西洋人は基本的に他人は自分とは異なる人間と見なす。言葉を尽くして自分の考えを伝え、他人に自分の考えを聞き取らせる必要があること。

西洋人は自分と異なる人間を自分と見なして考えを聞き取るために、言葉を尽くして自分の考えを伝えようとする試みを行うこと。

合計点: 12/16

2/2, 3/5, 3/3, 4/6, 0/1, 0/1

西洋人は自分と異なる人間を自分と見なして考えを聞き取るために、言葉を尽くして自分の考えを伝えようとする試みを行うこと。

西洋人は自分と異なる人間を自分と見なして考えを聞き取るために、言葉を尽くして自分の考えを伝えようとする試みを行うこと。

ご説明の流れ

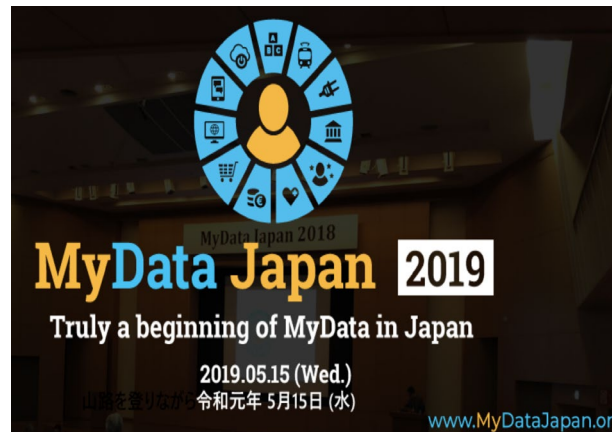


1. これまでの研究成果:

- A) 汎用基盤技術研究グループ(汎用G)
- B) 目的指向基盤技術研究グループ(目的G)
- C) 社会における人工知能研究グループ(社会G)

2. 今後の方向性

- **背景**: AI時代に即した社会制度の設計が必要
- **目標**: 社会系と技術系の研究者が協力してAIの社会的影響の分析を行い, 国際的に情報発信:
 - 国内外におけるAI倫理指針の策定
 - 法制度に関する提言
 - 日本に適した個人データ管理モデルの提案・実装
 - AI技術の公平性・信頼性等の規準策定・実装



成果3-1: AI倫理指針策定への貢献

- 人工知能学会倫理指針 (2017)
- 総務省AIネットワーク社会推進委員会
 - AI開発ガイドライン: **OECD**に提案 (2017)
 - AI利活用ガイドライン (2019)
- 内閣府人間中心のAI社会原則 (2019)
 - AI ready な社会の在り方: **G20**に提案
- **IEEE Ethically Aligned Design (2019)**
 - **世界最大規模の電気電子学会の倫理指針**



<https://www5.cao.go.jp/keizai-shimon/kaigi/special/reform/wg7/20191101/shiryou1.pdf>



#IEEEprinciples2practice
<https://ethicsinaction.ieee.org/>

成果3-2: 個人情報共有システム

■ 個人情報の収集・管理:

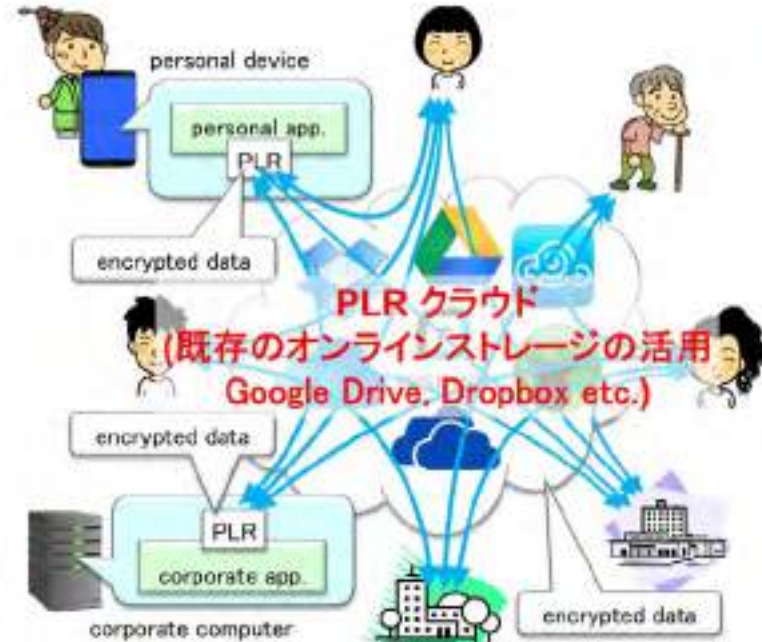
- **米国**: 企業主体, **欧州**: 国家で規制, **中国**: 政府主導

■ AIPでは**個人主体**のデータ共有システムを推進:

- データ持ち主の同意によるデータ公開
- 情報漏洩は起こらない
- 低コストで展開可能

■ **実用化**:

- 埼玉県立高の生徒12万人が2020年度**電子調査書**を実運用
- 今後の展開: 地域医療連携, 個人健康情報管理, 遠隔医療・看護, 行政手続き電子化など



ご説明の流れ



1. これまでの研究成果:

- A) 汎用基盤技術研究グループ(汎用G)
- B) 目的指向基盤技術研究グループ(目的G)
- C) 社会における人工知能研究グループ(社会G)

2. 今後の方向性

機械学習の国際研究拠点へ

19

- 基礎から応用・社会まで一貫通貫の研究体制
- 産学官で連携して、研究成果を国際的に発信
- 国際的な高度AI人材の登竜門となる
- **汎用G**: 深層学習理論と深層学習超え技術を開発
 - **2015年以前**: 日本全体で、主要国際会議での論文採択はひと桁、オーラル発表はほとんどなし
 - **現在**: AIPからICML2020に18件, NeurIPS2020に21件 (うちオーラル2件, スポットライト7件), AISTATS2021に17件, ICLR2021に8件 (うちオーラル2件)
- **目的G**: 進行中の研究の社会実装を加速
 - 医療, 自然災害, 材料, 教育
- **社会G**: 社会規範の策定から実装にシフト
 - 政府統計, 個人情報管理, セキュリティ, 公平性