

# 理研における AIPプロジェクト の実施について



理化学研究所 革新知能統合研究センター  
センター長(内定者) 杉山 将



# 革新知能統合研究センターについて 2

- 理化学研究所は、文部科学省が進めるAIP（人工知能/ビッグデータ/IoT/サイバーセキュリティ統合プロジェクト）の研究開発拠点として、4月14日付で**革新知能統合研究センター**を設置

- センター長（内定）：

**杉山 将（すぎやま まさし）**

東京大学大学院教授

※ 正式に着任するまでの間、  
理事 松本洋一郎がセンター長を兼務

- 特別顧問（内定）：

**金出 武雄（かなで たけお）**

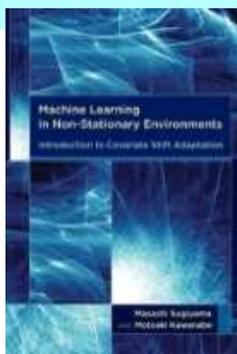
米カーネギーメロン大学教授



# これまでの研究： 機械学習の理論と応用

**教師付き学習**：人間が教師となり、  
コンピュータを学習させる

脳波によるコンピュータの操作  
(独Fraunhofer研究所との共同研究)

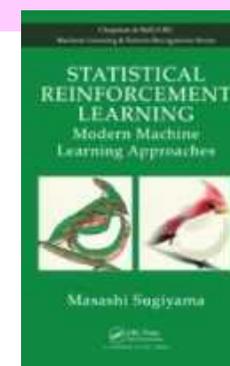
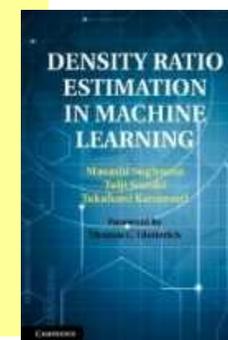


**強化学習**：エージェントが試行  
錯誤を通じて学習する

ヒューマノイドロボットの運動制御  
(NICT・ATRとの共同研究)

**教師なし学習**：コンピュータが  
人間の手を介さずに学習する

- データの可視化
- クラスタリング
- 密度比推定



## 非定常環境適応学習の手法と応用

- 変化する環境に適応できる新しい学習法を開発
- MIT Pressより**英語専門書を2冊出版(1冊は編者)**
- 主要文献の引用数は1000回以上(Google Scholar)
- 米国IBM社Faculty Awardなど**4件の受賞**

### ■ 応用例:

- 顔画像からの年齢推定(NEC)
- ヘルスケアデータの解析(NEC)
- 会話からの話者識別(ヤマハ)
- 日本語文章の単語分割(IBM)
- ブレイン・マシン・インターフェース(独Fraunhofer研究所)
- 半導体ウェハー製造装置の高速化(ニコン)
- 加速度データからの行動認識(NTT)
- ロボットアーム制御(独Max-Planck研究所)



## 強化学習のアルゴリズムと応用

- スパースデータから精度良く学習できる独自のアルゴリズム群の開発
- 著書2冊
- コンピュータアートへの応用がMIT Technological Reviewに注目研究として取り上げられた



## 密度比推定:新しいデータ解析パラダイム

- 異常検出, 変化検知, 特徴選択, 因果推論など様々なデータ解析タスクを, 統一的な枠組みで高精度・高速に解決
- **情報処理学会長尾真記念特別賞, 文部科学大臣賞若手科学者賞, 船井情報科学振興財団船井学術賞など7件の受賞**
- Cambridge University Pressより**英語専門書を出版**



### ■ 応用例:

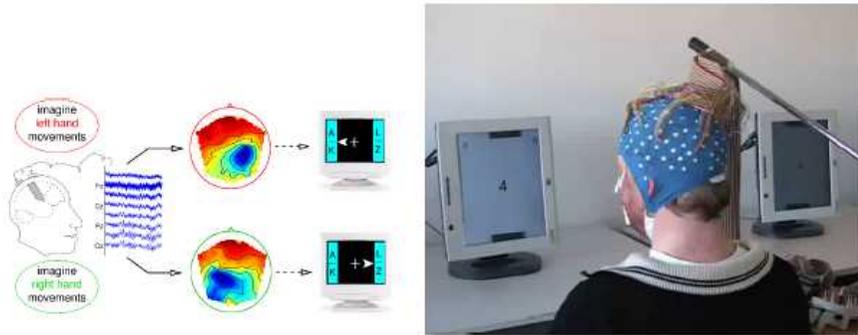
- 医療画像処理システム(富士フイルム)
- ヒューマノイドロボット制御(ATR), 脳波解析(トヨタ自動車)
- 製鉄プロセスの異常検知(JFE)・可視化(新日鉄住金)
- 光学機器の自動診断(キヤノン), ハードディスクの故障診断(IBM)
- 画像中の注目領域の抽出, 動画からのイベント検出(キヤノン)
- 半導体の自動検査(東レ), 半導体露光装置の異常検知(ニコン)

## 次元削減のアルゴリズム

- 古典的なフィッシャー判別分析を, アルゴリズムの簡便さを維持したまま, データ空間の局所的な構造を取り込めるように改良
- 主要論文の引用数は1000回以上(Google Scholar)

# ブレイン・コンピュータインターフェース

- 脳波(EEG)でコンピュータを動かす
  - 手足が動かない患者でもコンピュータが使える
- 少数のデータで精度の良い脳波認識を実現



ドイツ・フラウンホーファー研究所との共同研究

## 画像・動画処理

- 画像からの注目物体検出
  -
- 動画からのイベント検出
  -
- 事前の学習は全く必要なく、与えられた画像や動画から即座に検出！

キヤノンとの共同研究

## 顔画像からの年齢認識

- 実験室と実環境での照明条件, カメラ設置条件, 顔向きの違い, 人種の違いなどを克服

実験室 (室内照明環境・正面顔・・・)

実環境 (応用現場の環境) (日光/暗めの照明環境・横向き顔・・・)

NECソリューションイノベータとの共同研究

## コンピュータ・アート

- 「筆ロボット」が写真を毛筆画風に自動変換





## ヒューマノイド・ロボット制御

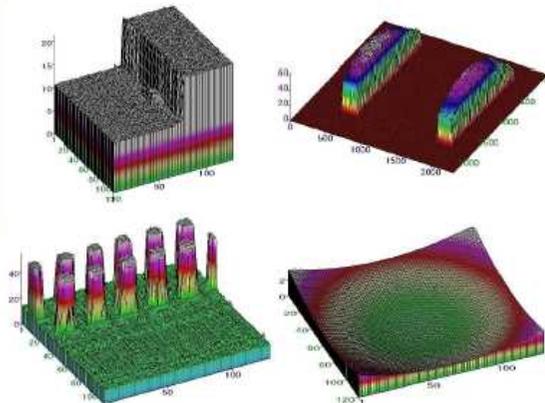
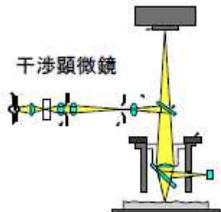
- 人間のような多自由度を持つロボットを、わずかな学習で正確に制御



ATR, NICTとの共同研究

## 微細表面形状測定

- ナノメートル単位の形状を光干渉を用いて一瞬で測定
- 局所的な類似性を用いて、高精度に推定



東レエンジニアリングとの共同研究

## 日本語単語分割

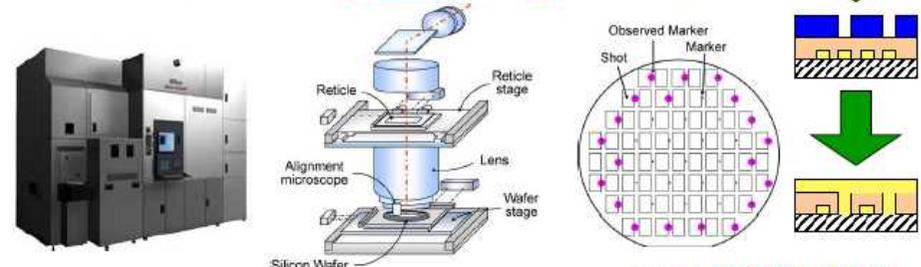
- 適応元データ: 会話辞典の例文
  - (例)こんな／失敗／は／ご／愛敬／だ／よ／.
- 適応先データ: 医療マニュアル
  - (例)細胞膜には受容体があり、これによって細胞を識別することができます.
- 異なる分野への**適応**を実現

手法	従来法	提案法	適応先ラベル有
F値	92.30	94.46	94.43

IBMとの共同研究

## 半導体露光装置におけるウェハの位置合わせ

- 近年のシリコンウェハは**多層構造**を持つ
- 何度も回路パターンを露光するため、**正確な位置合わせ**が非常に重要
- 位置合わせに**最適なマーカー**を選択



ニコンとの共同研究

# 人工知能研究の位置づけ

- 工学や科学の難問を抽象化し、  
数学とコンピュータを駆使して問題解決！

## 工学応用

テキスト, 音声, 画像, 映像,  
ロボット, 広告, 電子商取引

## 科学応用

医療, 生命, 物理, 化学,  
材料, 宇宙, 地学, 社会

# 人工知能

## 数学

ベクトル, 行列, 微分, 積分,  
確率, 統計, 論理, 最適化, 幾何

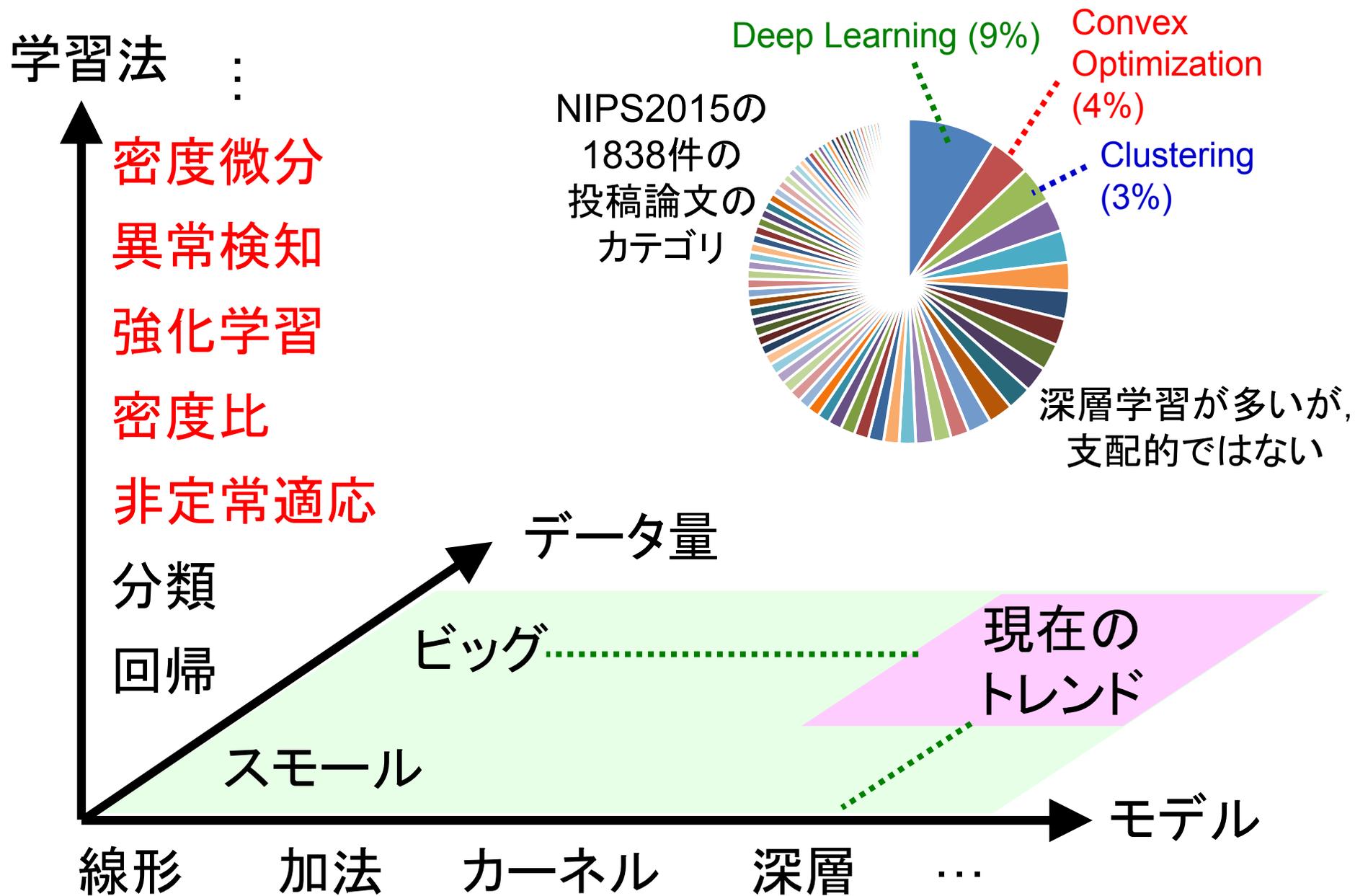
## コンピュータ

プログラミング, ハードウェア,  
ネットワーク

- **欧米**: 巨大民間企業等が数百億～数兆円規模の莫大な予算を投じて研究開発を開始
  - Google, Microsoft, Facebook, Amazon, Toyota, OpenAI,...
  
- **国内**: 政府が中心となって, 数十億～数百億円規模の予算を幅広い分野に配分

# 国際会議NIPS2015での アジア勢の動向

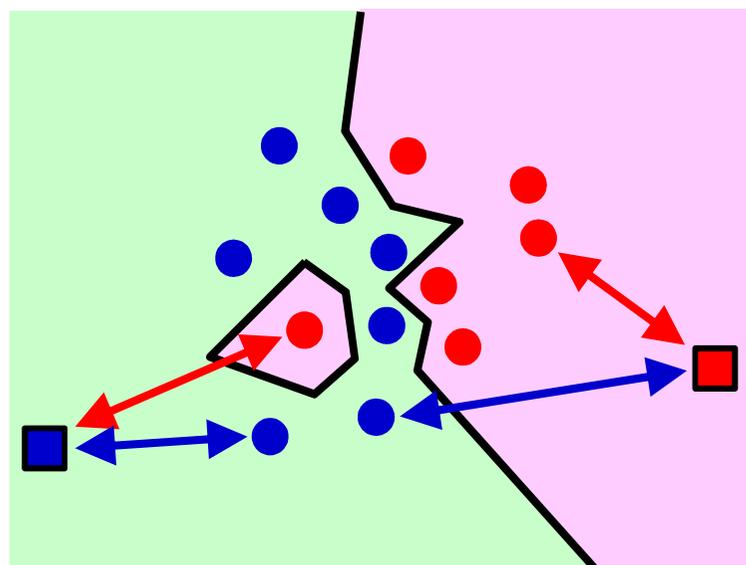
- 著者のメールアドレスに.jpが含まれる採択件数は**403件中8件(約2%)**:
  - .cnは13件, .krは0件. ただし, 中国・韓国は欧米の大学・企業に所属している学生・研究者が多数いる
- 体感的には, NIPSでの**日本人の存在感は皆無**:
  - 韓国は日本よりやや存在感があり,  
中国はかなり溶け込んでいるという印象
- しかし, 現在までの日本の情報科学の幅広い分野での研究の蓄積, 研究者の質の高さを鑑みれば, 今後飛躍をするチャンスは多いにある!



# ビッグデータ + 深層学習 = 究極の人工知能？

- データがたくさんあれば, 古典的な**最近傍分類**でもそこそこうまくいく

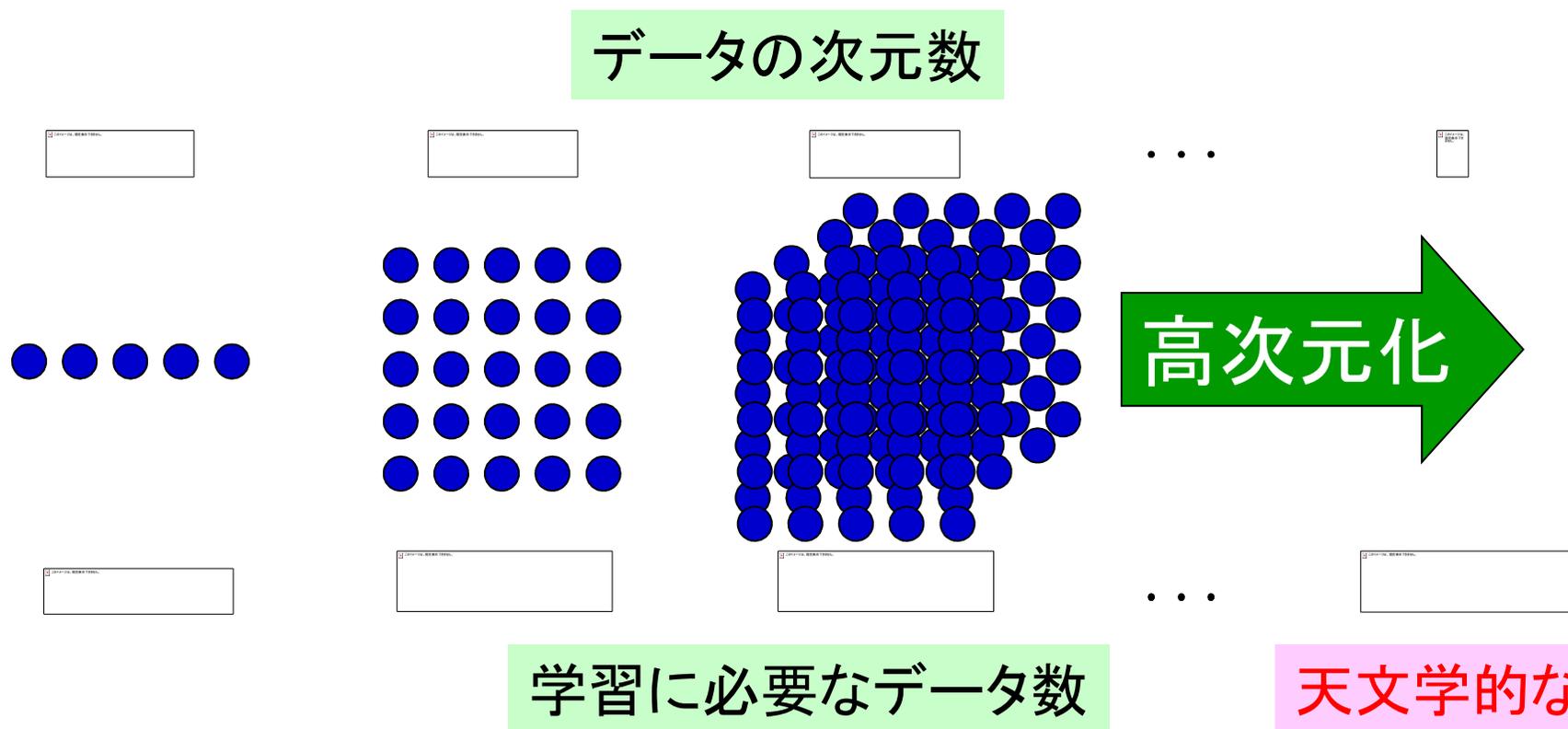
データ	深層学習 (CNN)	最近傍分類
印刷数字60万文字	0.01%	0%
手書き数字80万文字	0.11%	0.97%



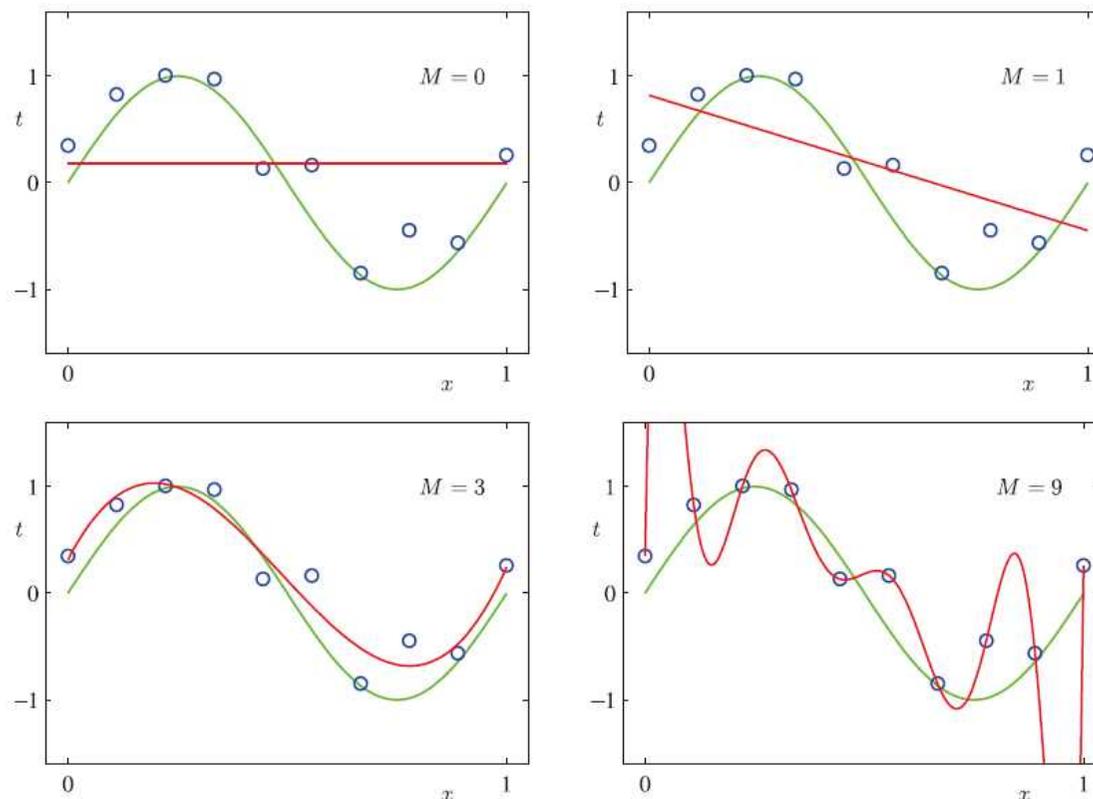
[PRMU2016/2内田ほか]

**最近傍分類:**  
一番近くのデータと  
同じクラスに分類

- センサーの数を増やすとデータの次元数が増加
  - 学習に必要なデータ数  $\square$  は, データの次元数  $\square$  に対して**指数関数的に増加**
- いくらビッグデータを集めても, **データはまばら**



- **汎化能力**: 教わっていないことを、過去の事例から類推する能力
- まばらなデータの間を補完・予測

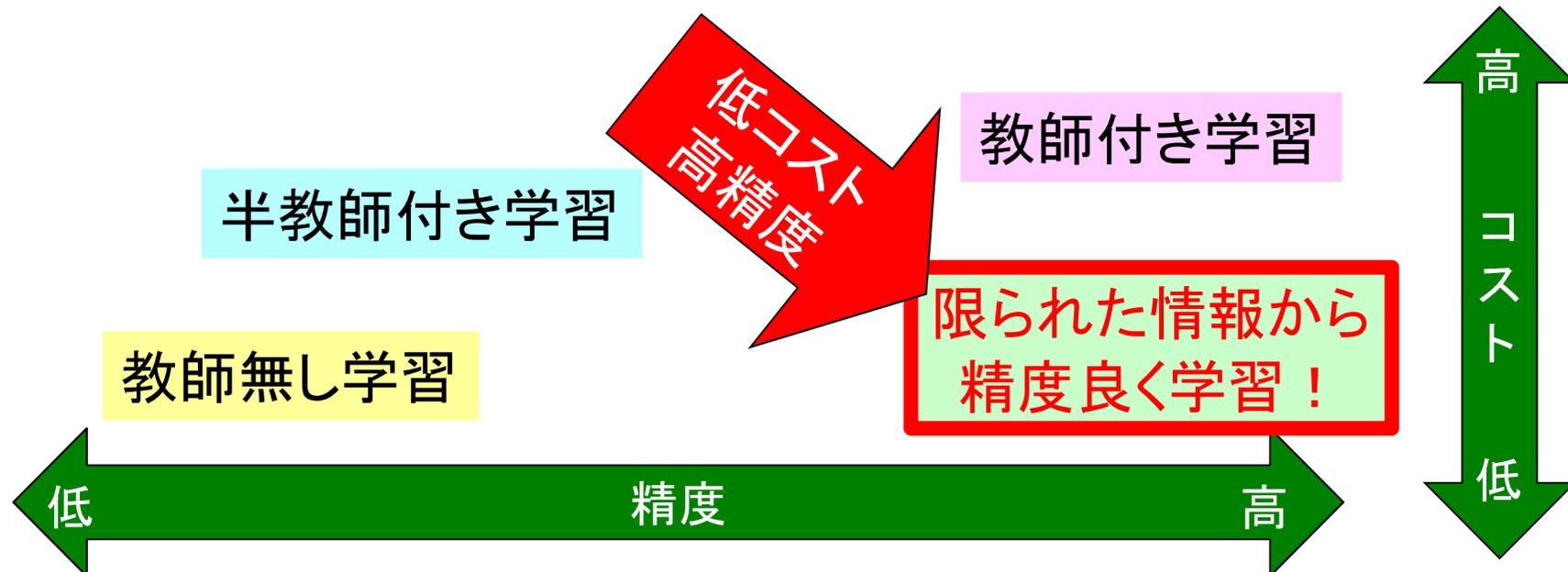


[パターン認識と機械学習 Bishop2006]



- 6面体のさいころを何回も振れば、その平均値は大体3.5
- 真の期待値(答え)がわからなくても、**たくさんのデータの平均**を取れば大体答えがわかる
- **中心極限定理**: **同じさいころ**を  $n$  回振ると、その平均値は本当の期待値に  $1/\sqrt{n}$  の速さで近づく
- ビッグデータを用いれば答えが簡単に予測できそう？

- 同じさいころを多数回振れない
  - 例: 同じ病状の患者はそれほどたくさんいない
- そもそも答えのないデータばかり
  - 例: ドローンで橋梁の画像をたくさん撮っても、危険個所は人間が特定しないといけない
- 限られた情報からの学習が重要！



# AIP: Advanced Integrated Intelligence Platform Project

## 人工知能/ビッグデータ/IoT/サイバーセキュリティ統合プロジェクト

平成28年度予算額 : 5,448百万円(新規)  
(関連する既存事業(2,849百万円)を含む)  
※運営費交付金中の推計額含む

### 【国際的な動向】

- 各分野でのビッグデータの集積、センサーの量的・質的拡大(IoT: Internet of Things)
- 人工知能に50年来の大きな技術的ブレークスルー(自ら特徴を捉え進化する人工知能が視野)
- 一方、高度化する脅威に対するサイバーセキュリティの確保(ますます巧妙化しており、人材育成が必須)

### 【文部科学省の対応】



- (1) 文部科学省が持つビッグデータの解析(コホート、環境のデータなど多様)を通じて、新たな価値を創造。
- (2) そのため、革新的な人工知能技術を開発・活用
- (3) ビッグデータの充実のため、高度なセンサー/IoT技術を活用。あわせて、堅牢なセキュリティを構築。

〔経済産業省・総務省との連携を呼びかけ、基礎研究から社会応用まで、一体的に実施する体制を構築〕

### AIPセンター(理化学研究所)1,450百万円

- I. 人間の知的活動の原理に学んだ革新的な人工知能の基盤技術を開発。
- II. 人工知能とビッグデータにより複数分野においてサイエンスを飛躍的に発達させる。
- III. 具体的な社会・経済価値を創造する多数の応用領域の社会実装に貢献。
- IV. 人工知能等が浸透する社会での倫理的・社会的課題等に対応。
- V. データサイエンティスト、サイバーセキュリティ人材等を育成。

### 戦略的創造研究推進事業(一部)(科学技術振興機構)

新規採択課題分 1,150百万円  
関連する既存採択課題分 2,849百万円  
※運営費交付金中の推計額

- 大学等の研究者から広く提案を募り、組織・分野の枠を超えた時限的な研究体制を構築して、**戦略的な基礎研究を推進**。



### 【目指す成果】

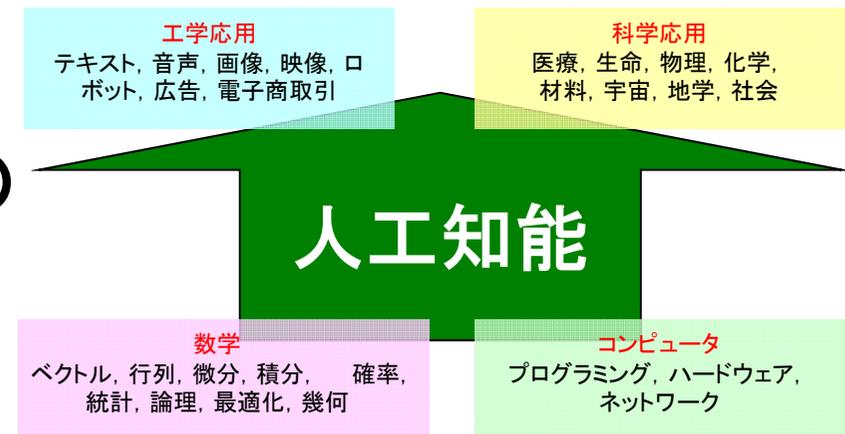
人々と社会のための知能とイノベーションの創出に向けて、世界的に優れた競争力を持つ研究者を結集、最先端研究を統合。我が国が直面する労働力減少、高齢化社会の中でも、

- ・生産性の大幅な向上による**経済成長への貢献**、
- ・一人ひとりに優しい**社会構築**(医療・介護等)を実現

# 理研AIP事業で行う 研究開発の考え方

- 数理学等に立脚した人工知能の原理・原則の  
解明を進め、世界をリードする**革新的人工知能**  
**基盤技術を構築**

- サイエンスや実社会などの  
幅広い出口に向けた  
**応用研究を進める**



- 未来の科学研究に必要となる**データ構造, データ**  
**取得技術の世界標準化を図る**

## ■ 超高齢社会へ向けた医療サポート

- 動画認識・センサ情報解析・アクチュエータ制御技術を融合し、高度な手術を支援
- 機械学習による電子医療記録の自動解析・予後予測
- 会話を通じて高齢者の認知機能を維持向上

## ■ 老朽化が進むインフラへの対応

- 自動操縦ドローンを駆使し、動画や打診音から橋梁などの危険個所を自律的に特定

## ■ 甚大な自然災害への対応

- ビッグデータ解析技術とシミュレーション技術を統合し、甚大災害をもたらす異常気象を適確に予測し、被害を最小限に抑え迅速に復旧できる社会システムを構築

## ■ プライバシー

- プライバシー, 説明責任, 公平性, 差別配慮, セキュリティなどを考慮した人工知能技術の開発

## ■ 細分化が進む科学研究への対応

- 論文・特許・実験結果をもとに科学研究の発展を支援
- これまで埋もれていた発見を見逃さない技術を開発
- 次に実験すべき項目を過去のデータに基づいて決定
- マテリアルズ・インフォマティクスなどへの応用

## ■ 倫理

- 人工知能技術が社会に浸透する際のELSI問題や社会的影響を, 人文科学・社会科学の視点で検討

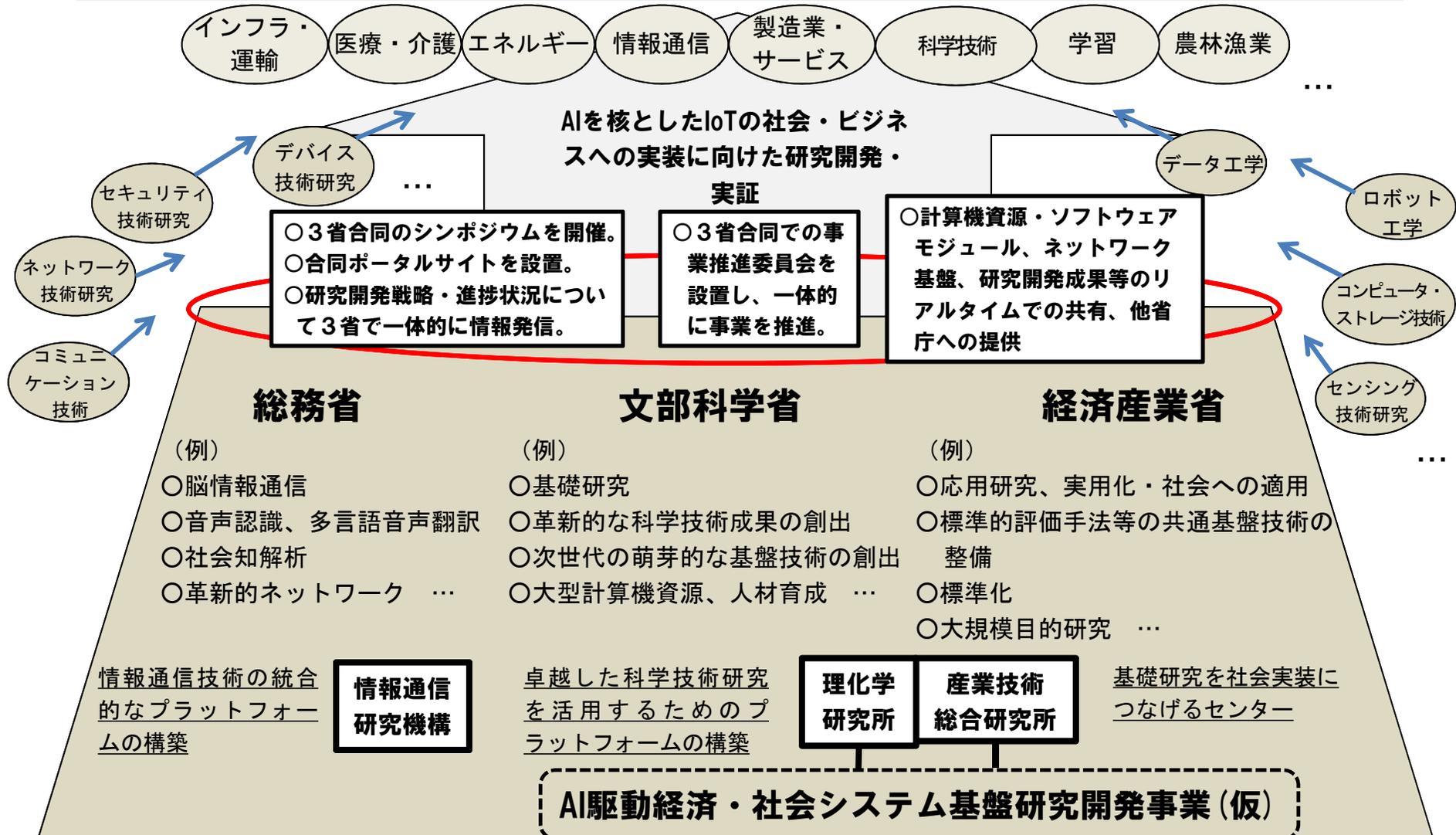


# 理研AIP事業で行う研究開発の例 <sup>20</sup>

- データサイエンティスト, サイバーセキュリティ人材等の育成
  - クラウドソーシングを用いた遠隔データ解析教育
- 革新知能プラットフォーム構築
  - 深層学習等のビッグデータを用いた学習の革新
  - 疎・不完全・超高次元データからの高精度学習の実現
  - ストリーミングデータに対するリアルタイム学習の実現
  - 学習アルゴリズムの選択・調整を自動化
  - CPU, GPU, ディスクI/O等を考慮した学習技術の実現
  - 最適なデータ収集戦略の策定

# 次世代の人工知能技術の研究開発における3省連携体制

- (1) 各分野でのビッグデータの集積、センサーの量的・質的拡大 (IoT: Internet of Things)。
- (2) 人工知能の50年来の大きな技術的ブレークスルー (自ら特徴を捉え進化する人工知能を視野)。
- (3) 3省連携による研究開発成果を関係省庁にも提供し、政府全体として更なる新産業・イノベーション創出や国際競争力強化を牽引。



# All Japanを超えた 国際研究拠点を目指して

- 欧米の大学・企業は「外国人」が集う国際拠点
  - 助っ人外国人を雇うのではなく、国際拠点を目指す
  - 外国人率30%以上を目指す
- Money: That's What I Want ?
  - 国際的に競争力のある待遇で有力な研究員を雇用
- 欧米と比べて、日本の大学では博士課程に進学する学生が非常に少ない
  - 修士・博士課程の学生のサポートを充実
  - 研究員に長期のポストを提供
  - 企業・大学と連携し、人材の流動性を高める

# 日本は欧米の周回遅れ? → No! 23

- 欧米は欧米のやり方で山頂を目指している
- その道を日本が今から追いかけても、  
追い越すことは難しい
- 日本は日本のやり方で  
山頂を目指すべき!

