

IoT時代の機械学習技術

～環境知能の実現にむけて～

理化学研究所 革新知能統合研究センター

NTT コミュニケーション科学基礎研究所

上田修功

Naonori Ueda



自己紹介（最近の研究プロジェクト）

2013.4～現在：
NTT 機械学習・データ科学センター
（センター代表）

2016.4～現在：
NTT コミュニケーション基礎研究所
上田特別研究室長(NTTフェロー)

2016.9～現在：
理化学研究所
革新知能統合研究センター
（副センター長）
防災科学チームリーダー
iPS細胞連携医療リスク回避チームリーダー

2010.3～2014.2：
喜連川FIRST

2015.10～2021.3：
JST CREST
統計計算宇宙物理

2014.7～2016.3：
NICT Smart City

2016.4～2026.3：
Riken AIP

講演内容

Part I: 人工知能&機械学習

Part II: 環境知能とその応用

Part III: 理研AIPのご紹介

講演内容

Part I: 人工知能&機械学習

Part II: 環境知能とその応用

Part III: 理研AIPのご紹介

人工知能と機械学習

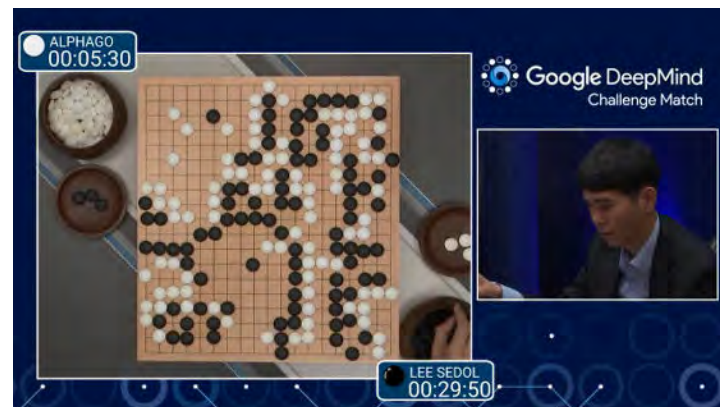
- 自動運転車，会話ロボット，コンピュータ囲碁など，私達の身の回りの様々な場面で「人工知能」が利用されるようになってきた。



<http://www.cnn.com/2015/10/14/tesla-rolls-out-autopilot-technology.html>



http://www.softbank.jp/corp/group/sbr/news/press/2014/20141029_01/



<http://gigazine.net/news/20160317-google-alphago/>

- 「人工知能」の背後では，コンピュータにヒトの ような学習能力を持たせる**機械学習**の技術が用いられている。

・1950～80年：人工知能研究の開花

数値計算から記号処理へ（人間の知能獲得の模倣）

・1980年台：人工知能研究の産業化

エキスパートシステム、ニューラルネットが大流行

・1990年台：データマイニング応用

計算機能力の飛躍的な向上

1997年：DeepBlueがチェスチャンピオンG.Kasparovに勝利

・2000年～：統計的機械学習

テキスト、SNS、関係データなどの非構造化データからの知識獲得、レコメンデーションなどのデータマイニング

・2010年～：ビッグデータ、最適化

ワトソン（ジェパティ）、Deep ニューラルネットワーク（Googleの猫）

機械学習研究の有用性

機械学習技術

数理統計・最適化理論に基づく汎用データ解析技術

(汎用性：パターン認識、言語処理、ロボット工学、BMI, バイオインフォマティクス、データマイニングなど広範囲の分野で利用されている土台技術)

ビッグデータ時代では、厳密な物理モデルだけでは解析できない対象・問題が主流



「物理モデル」から「目的達成のモデル化」へ

サービス科学(サービスのための方法論の科学)
の土台技術として機械学習技術が有望

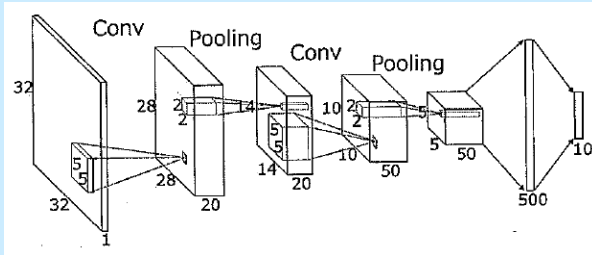
機械学習技術の特徴

- ・汎用性

応用分野に依存しない手法
(多くの応用分野で要素技術を提供)

- ・工学/数理アプローチ

真理の解明ではなく、十分条件として
データ生成メカニズムを理論的に説明で
きればよい



0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

印刷文字

約60万データで学習：
認識率 99.99%

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

手書き文字

約80万データで学習：
認識率 99.89%

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

マルチフォント文字

約6000種類のフォント：
認識率 96.4%

出展：内田他：信学技報、2016

- DNN(CNN)は膨大な学習データで極度の非線形関数を学習している

これって全記憶!?

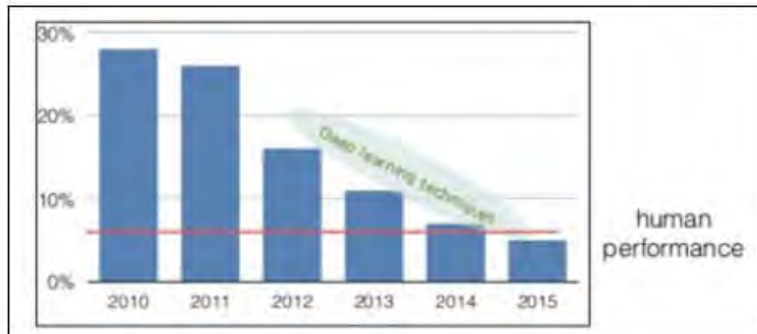
cf. ハミング距離でk-近傍法では印刷文字だと100%の精度。ただし、
一文字当たりの認識速度は約1000倍

- 切り出した画像を用いて学習

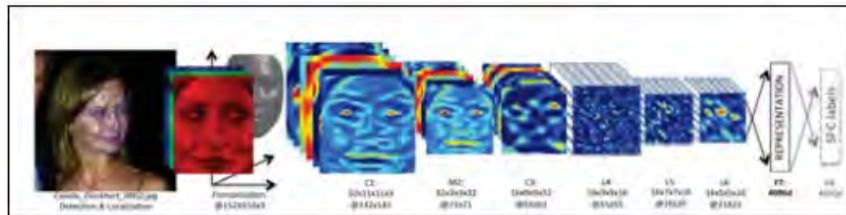
(文字や言語は有限なので学習データ収集は比較的容易)

深層学習の進化

人間の精度を超える!?



自然画像分類のエラー率
素人:5.1~15%
プロ:2%
ResNet (Microsoft):3.57%



顔識別の精度
人間:97.53%
DeepFace (Facebook):97.35%
FaceNet (Google)99.63%

プロ棋士 vs 囲碁AI (AlphaGo)
AlphaGo (Google DeepMind)が
4勝1敗(5戦中)

深層学習に関する知見

- ・観測データと教師信号が正確、かつ、**大量に与えられた教師有り学習**では現状では**最強**のツール
- ・画像、音声、自然言語処理では、当面、DNNは**必須要素技術**
- ・ただし、**学習データが容易に準備できない応用**（観測データと教師信号の対が不明確なケース）や、結果に対する**説明が必要な応用**（因果分析）ではDNNの適用は困難

科学研究の変遷

1. 実験科学

アリストテレスの天動説（自然哲学）

2. 理論科学

ニュートン（プリンキピア）

3. 計算科学

数値計算, シミュレーション

4. データ集約型科学（第4の科学）

データからの仮説発見

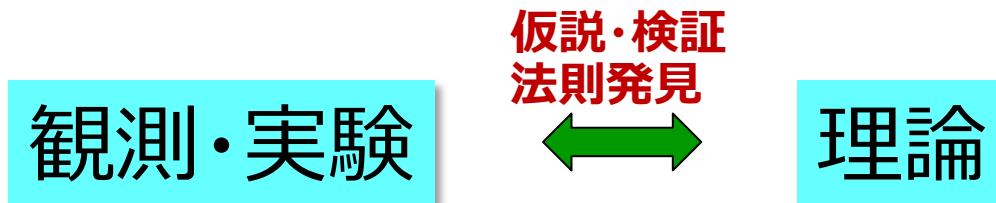


シミュレーション科学!?

シミュレーション科学

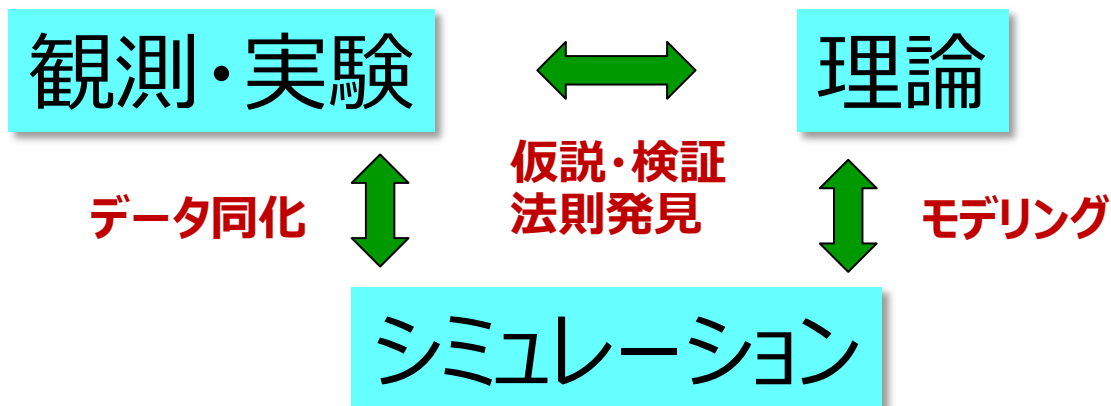
データ科学の新展開

伝統的な科学研究： 人手による仮説策定と修正の試行錯誤



シミュレーションデータ駆動型科学研究：

少数観測データから、シミュレーション技術により
ビッグデータを創出し、新たな知識を自動獲得



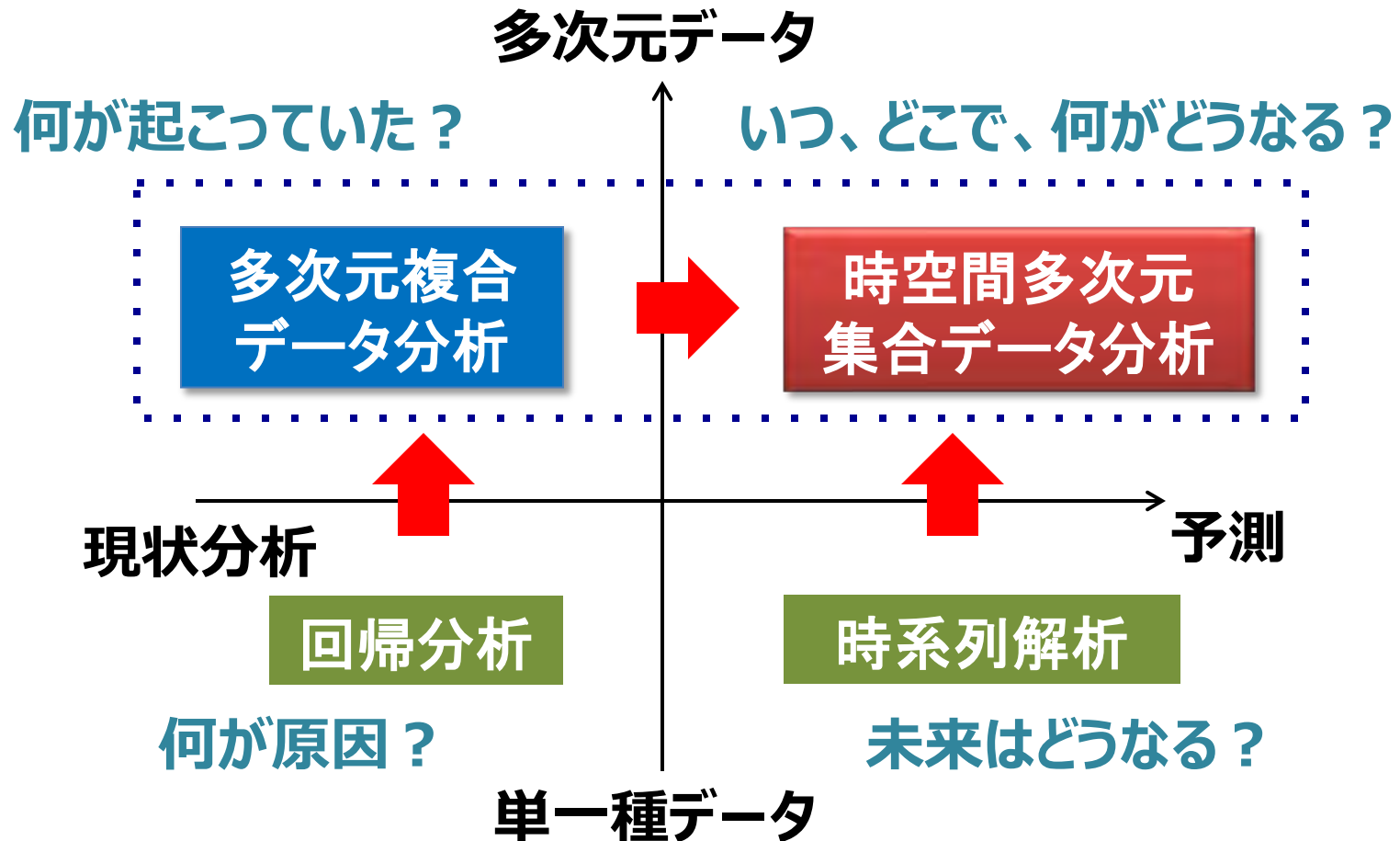
講演内容

Part I: 人工知能&機械学習

Part II: 環境知能とその応用

Part III: 理研AIPのご紹介

IoT時代での基盤分析技術



オープン・スマートシティを実現する ソーシャル・ビッグデータ利活用・還流基盤

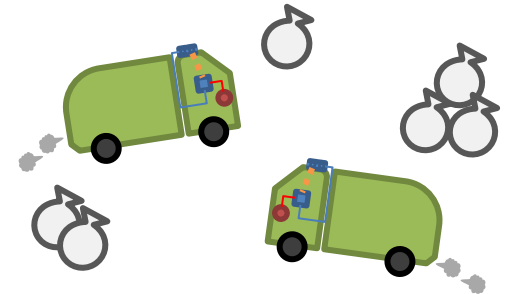


慶應義塾大学, 東京大学, 東京電機大学, 日本電信電話



背景

- ゴミは人の営みと切っても切り離せない。
- ゴミ収集は地方自治体の基本業務の一つ
- 将来的に安定してゴミ収集を運営するためには？
 - ゴミ収集の効率化
 - ゴミ減量



センサ搭載ゴミ収集車によるスマートゴミ収集

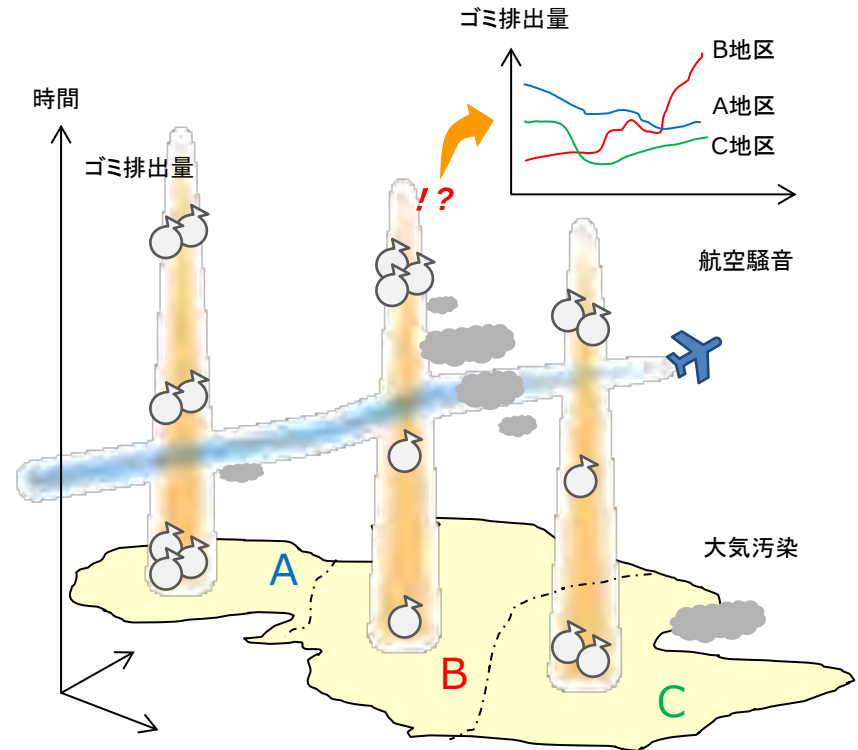
環境センサによる時空間都市イベント検出

• 環境データによって検出できる都市イベント

- 大気汚染ホットスポットの出現
- 騒音の発生
- ゴミ排出量の急激な増加
-

• 検出した都市イベントを活用したスマート都市運用

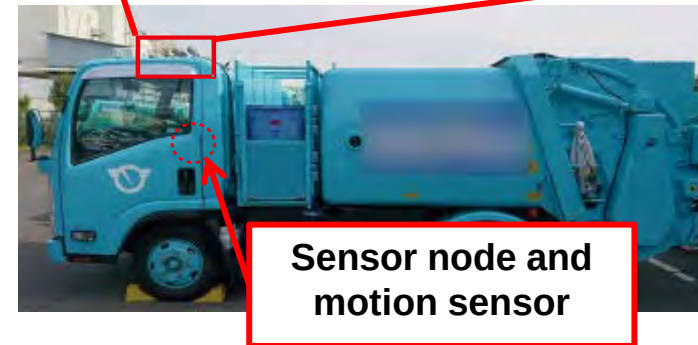
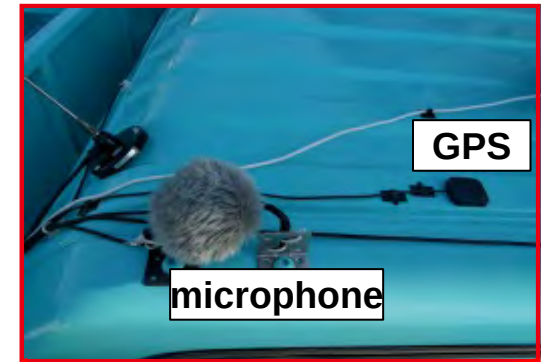
- 詳細な大気状態を市民に通知
- 騒音制御
- スマートなゴミ収集
 - 例, ゴミ収集経路の最適化
-



ゴミ収集車のモーション測定

• ゴミ量推定に利用したセンサ

- モーションセンサ
 - 3軸加速度, 3軸角速度
 - 100Hz
 - 車駆動中と回転板回転中の振動の変化を検出
- GPS
 - 1 record / sec
 - ゴミ収集車の位置を測定
- マイク (学習データ作成時のみ)
 - 耐水マイクを車外に設置
 - 音声を聞き, 回転板の動作をラベリング

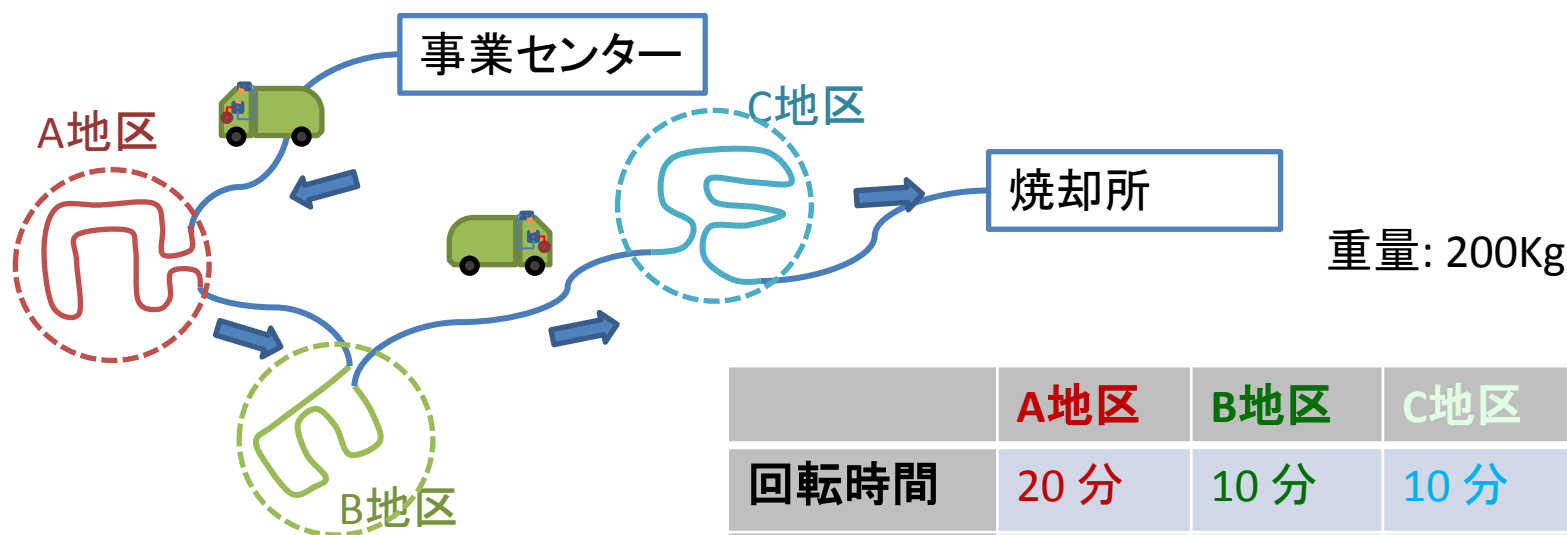


ゴミ量の推定

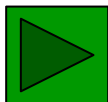
□ 仮定

■ ゴミの量に比例して回転板の動作時間が決まる.

□ 焼却所などで測定したゴミの重さを回転板の動作時間の長さで按分し, ゴミの量を推定

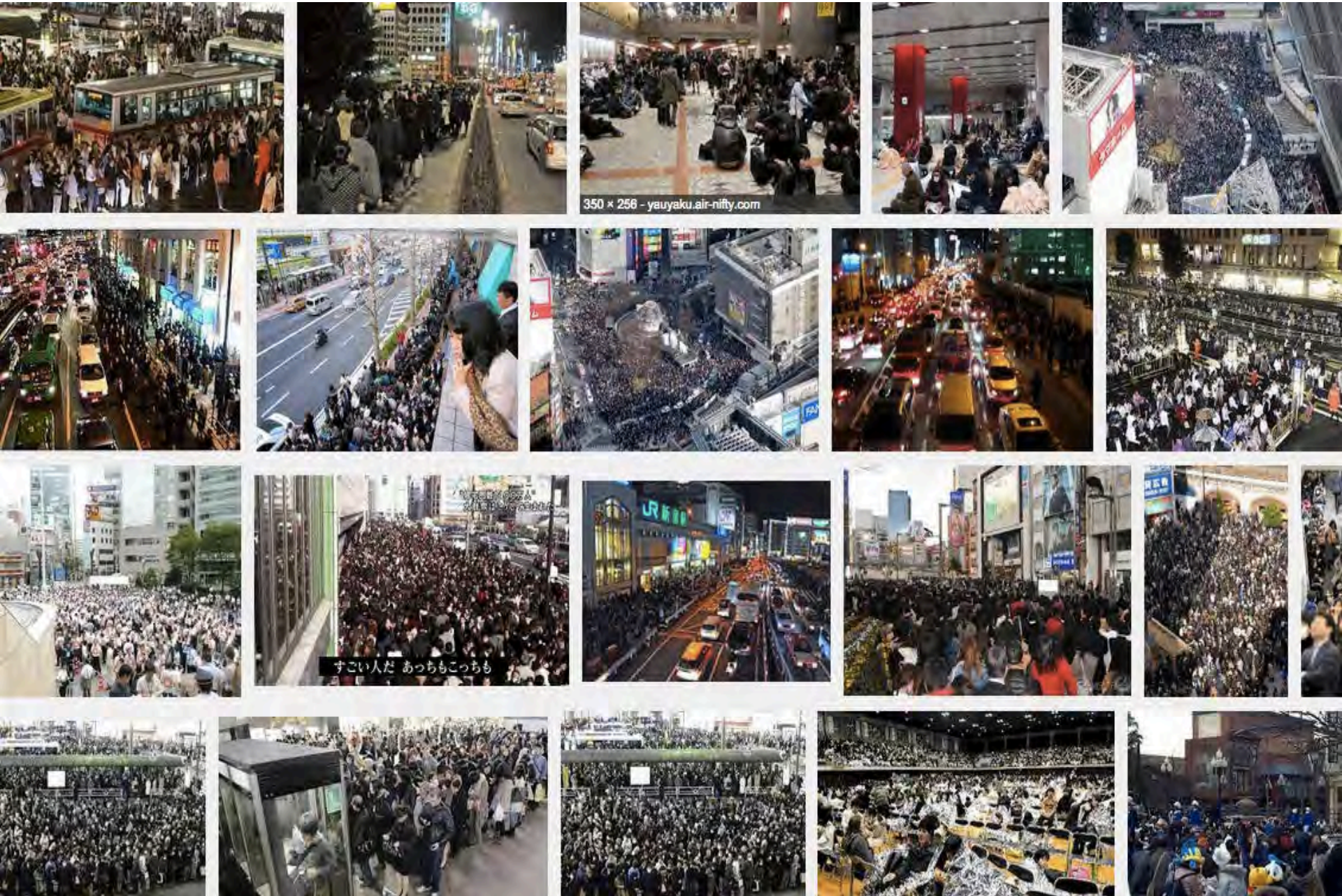


	A地区	B地区	C地区	合計
回転時間	20 分	10 分	10 分	40 分
推定ゴミ量	100Kg	50Kg	50Kg	200Kg



帰宅難民続出 (March 2011)

21



リアルタイム時空間予測・制御技術

ナビゲーションが満たすべき要件

1. 集団最適性

個人ではなく、**集団として最適**な誘導

2. リアルタイム性

現時点での状況(観測)に応じた**リアルタイム**な誘導

3. 先行性

渋滞が生じる前に、予測に基づいた**先行的**な誘導

マルチエージェントシミュレーション に基づく環境知能の実現

- 起源：**ユビキタスコンピューティング** (Weiser, 1991):
コンピュータシステム = ユーザ(人間) + **環境**
- エージェント：人間の知的活動を支援
(**IA: Intelligence Amplifier**)
- **マルチエージェント**：集合知(協調により $1 + 1 > 2$)
cf. アンサンブル学習
- Trillion Sensors Universe時代の到来
(**AI: Ambient Intelligence**)

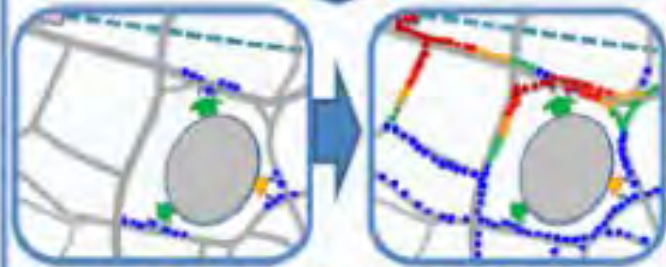
学習型マルチエージェントシミュレーション²⁴

従来型MAS*



パラメータ

誘導シナリオ



シミュレーション環境

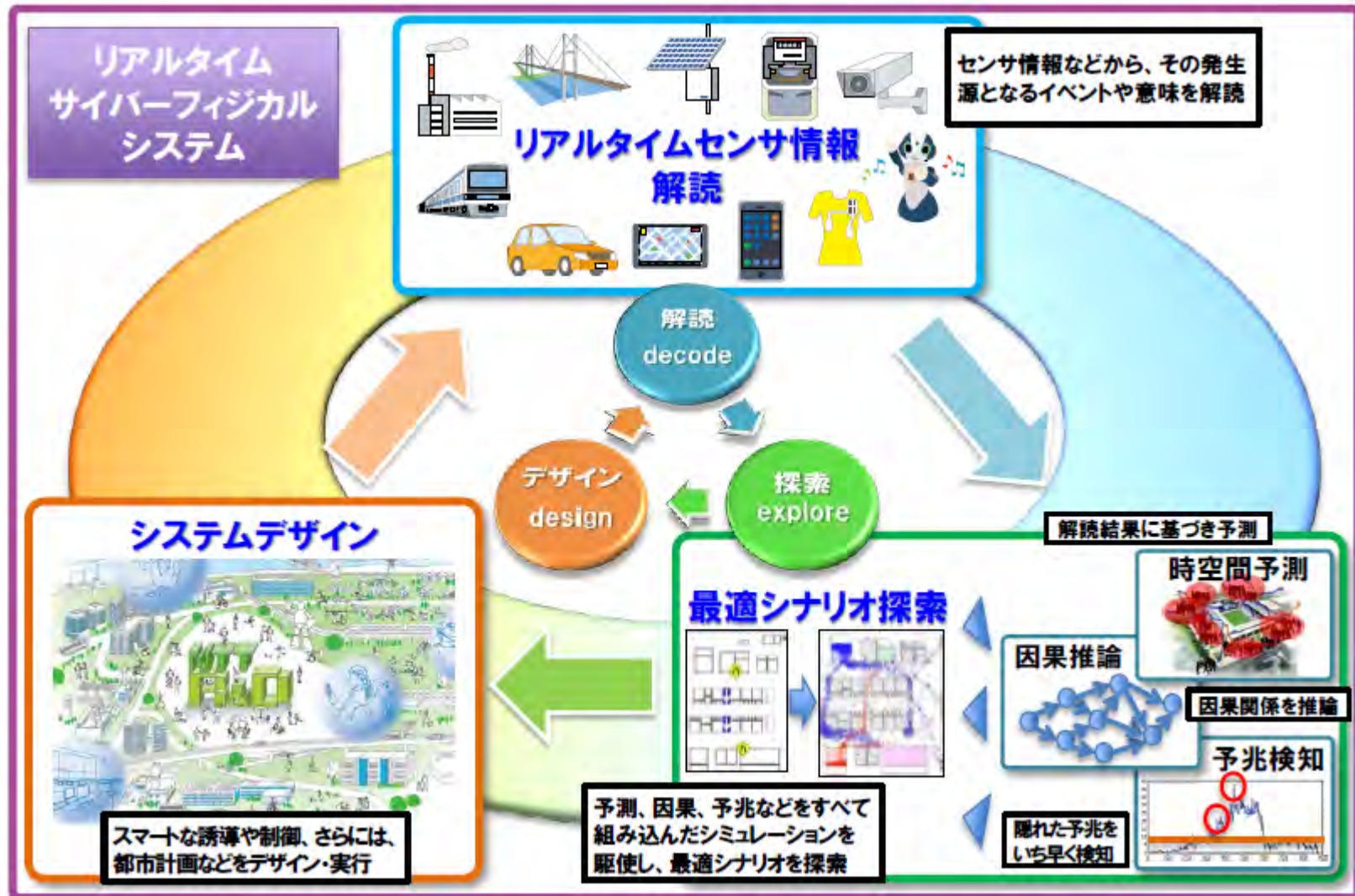
- MASのパラメータや誘導シナリオを手入力
- あらかじめ決められた誘導シナリオのオフライン評価

学習型MAS



- MASの**パラメータ**を観測データに基づき**自動学習**
- 事前の誘導・What-ifシナリオを用いた学習により、混雑リスクを回避可能な**最適誘導策**をオンラインで随時自動導出

多種多様なセンサ情報がNWに接続され 人間社会で活用されるICT社会



講演内容

Part I: 人工知能&機械学習

Part II: 環境知能とその応用

Part III: 理研AIPのご紹介

第5回「未来投資に向けた官民対話」 (H28.4.12)における安倍総理の発言

- 人工知能の研究開発目標と産業化のロードマップを，本年度中に策定します．
そのため，産学官の叡智を集め，縦割りを排した『**人工知能技術戦略会議**』を創設します



- － **議長**：安西 祐一郎(独立行政法人日本学術振興会 理事長)
- － **顧問**：久間 和生(内閣府総合科学技術・イノベーション会議常勤議員)
- － **構成員**：
 - ・ 内山田 竹志(日本経済団体連合会未来産業・技術委員会共同委員長)
 - ・ 小野寺 正(日本経済団体連合会未来産業・技術委員会共同委員長)
 - ・ 五神 真(国立大学法人東京大学総長)
 - ・ 西尾 章治郎(国立大学法人大阪大学総長)
 - ・ 坂内 正夫(国立研究開発法人情報通信研究機構理事長)
 - ・ 松本 紘(国立研究開発法人理化学研究所理事長)
 - ・ 中鉢 良治(国立研究開発法人産業技術総合研究所理事長)
 - ・ 濱口 道成(国立研究開発法人科学技術振興機構理事長)
 - ・ 古川 一夫(国立研究開発法人新エネルギー・産業技術総合開発機構理事長)

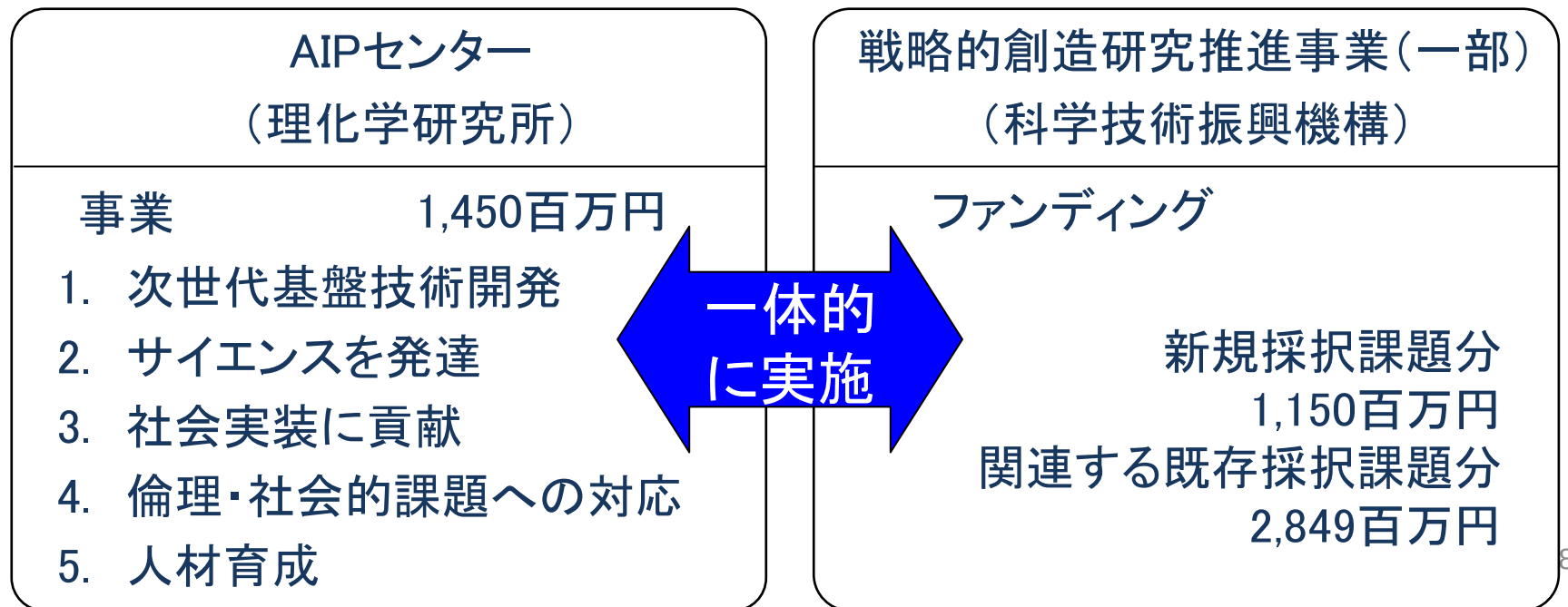
文科省AIPプロジェクト:

(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

- 理化学研究所は, 2016年4月14日付で
革新知能統合研究(AIP)センターを設置



AIPセンターの研究体制

29

世界を取り巻く広範な人工知能の応用分野
(企業・大学・研究所・理研内他センターなど)

目的指向基盤技術研究グループ:
実世界の複雑な問題を解決可能な形に抽象化

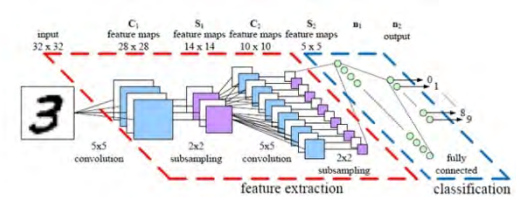
汎用基盤技術研究グループ:
抽象化された問題を解決する汎用技術の開発

社会における人工知能研究グループ:
人工知能の普及に伴う社会的影響を分析

AIPセンターの研究戦略1: 10年後を見据えた基礎研究を推進

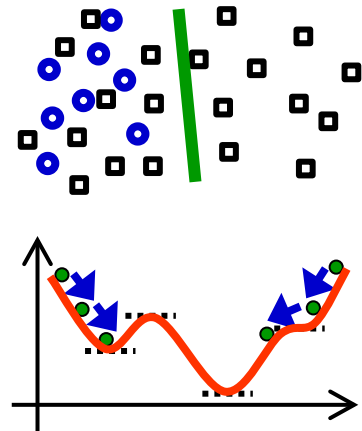
■ 主要な応用研究は予算規模の勝負になりつつあるが、**基礎研究は個人勝負**:

- AlphaGoはDeepMindの少数の研究者の成果
- 現在大流行している深層学習も、10年前は冷ややかに見られていた



■ **日本には優秀な理論研究者がいる**:

- **深層学習の原理**を理論的に解明し、更なる性能・効率の向上につなげる
- 深層学習で太刀打ちできない難題解決を目指した**次世代AI基盤技術**を開発する

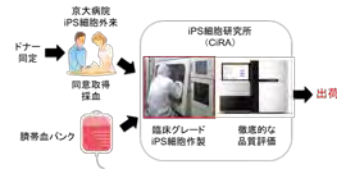


AIPセンターの研究戦略2:

応用研究は強力なパートナーと連携

■ 日本が強いサイエンス分野をAIで更に強化:

- 再生医療: 京都大学iPS細胞研究所
- モノづくり: 名古屋大学価値創造研究センター
- マテリアル: 物質・材料研究機構



■ 国内で取り組む必要のある社会的課題:

- 高齢者ヘルスケア: 国立がん研究センター, 東北メディカル・メガバンク機構
- 防災・減災: 防災科学技術研究所, 内閣府SIP防災
- 橋梁・トンネル等の検査: 内閣府SIPインフラ



AIPセンターの研究戦略3:

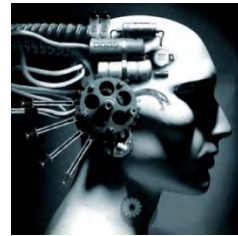
AIの社会的影響を分析

32

- 人工知能技術が社会に浸透する際の倫理的・法的・社会的影響や技術的特異点(シンギュラリティ)の影響の分析



- 人工知能の倫理的課題を数理的に解決:



- プライバシ: 暗号化データ上で動作する人工知能, 匿名化等によるデータ流通促進
- セキュリティ: 人工知能への敵対的攻撃に対する耐性
- 公平性: 予測・意思決定における公平性の保証

AIPセンターの運営戦略4: 人材育成を重視

33

- **現状**: 日本のAI研究者は層が薄い
- **高度な研究開発人材**を育成:
 - 大学の非常勤研究者経由で学生を集める
 - 連携企業の研究者・エンジニアを集める
- **データサイエンティスト**を育成:
 - 大学や研究所と協力して教育プログラムを実施

まとめ

パターン認識においてもDNNは全能の神ではない！

- ・深層学習でも適用困難な、**学習データが不完全**な教師有り学習の研究が重要

個体レベルの知能だけではなく、集合知、環境知能という視点も重要

- ・ミクロな構成要素が多数集まった**マクロレベルで創発**される「人工知能」研究も重要

ICTのパワーをenhanceし、日本を豊かにするAI研究であるべき

- ・**課題解決**（ex. ストレス社会でのQoL向上）を
目指した構成論的な研究アプローチが必要で、かつ、
出口戦略も重要

参考図書

KADOKAWA インターネット講座第7巻第5章

データに語らせる科学

～数式なしでわかる人工知能とビッグデータの新しい関係～
上田修功

第4の科学, IoT, 人工知能, 機械学習について, より俯瞰的な視点で, 時代と共にどのように生まれ進化してきたかを解説.