



「次世代人工知能・ロボット中核技術開発」
(人工知能分野) 中間成果発表会
－人間と相互理解できる人工知能に向けて－

次世代人工知能技術分野の狙いと全体像

平成29年 3月29日

国立研究開発法人 産業技術総合研究所
人工知能研究センター

辻井 潤一

国立研究開発法人 産業技術総合研究所

国立研究開発法人 新エネルギー・産業技術総合開発機構

NEDO委託事業「次世代人工知能・ロボット 中核技術開発」（次世代人工知能技術分野） 人間と相互理解できる 次世代人工知能技術の研究開発

研究開発項目①：大規模目的基礎研究・先端技術研究開発

研究開発項目②：次世代人工知能フレームワーク研究・
先進中核モジュール研究開発

研究開発項目③：次世代人工知能共通基盤技術研究開発

研究開発期間：2015年度～2019年度

先導研究期間：2015年度～2016年度（2015年度採択分）

2016年度～2017年度（2016年度採択分）

- 米国の巨大IT産業

- Internet 上のデータ、資金、研究者、開発者の集中
- 閉じたエコシステムの形成
- Start-Upの M&A

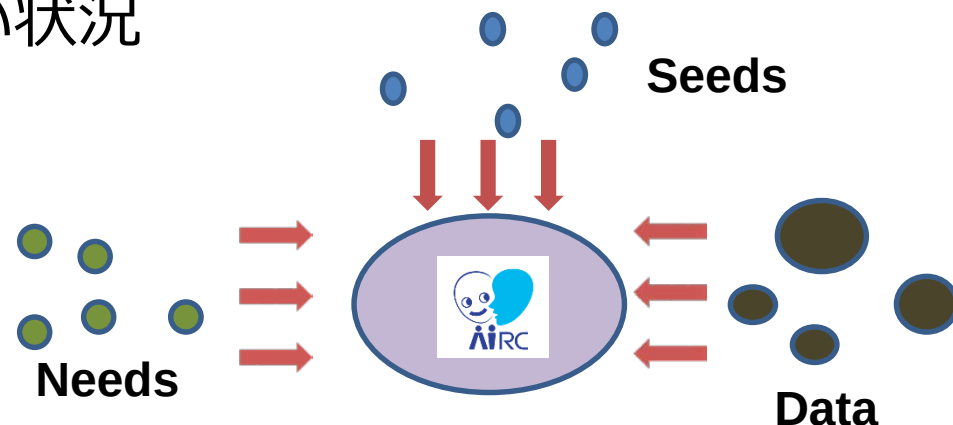
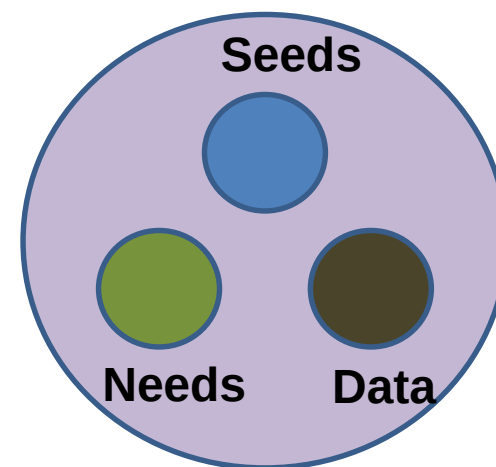
- 日本（ヨーロッパも）

- Needs, Seeds, Data のFragmentation
- 技術が社会に出てゆきにくい状況



- 開いたエコシステムの形成
- Open Innovation Hub

巨大IT産業(G,M,F,A)

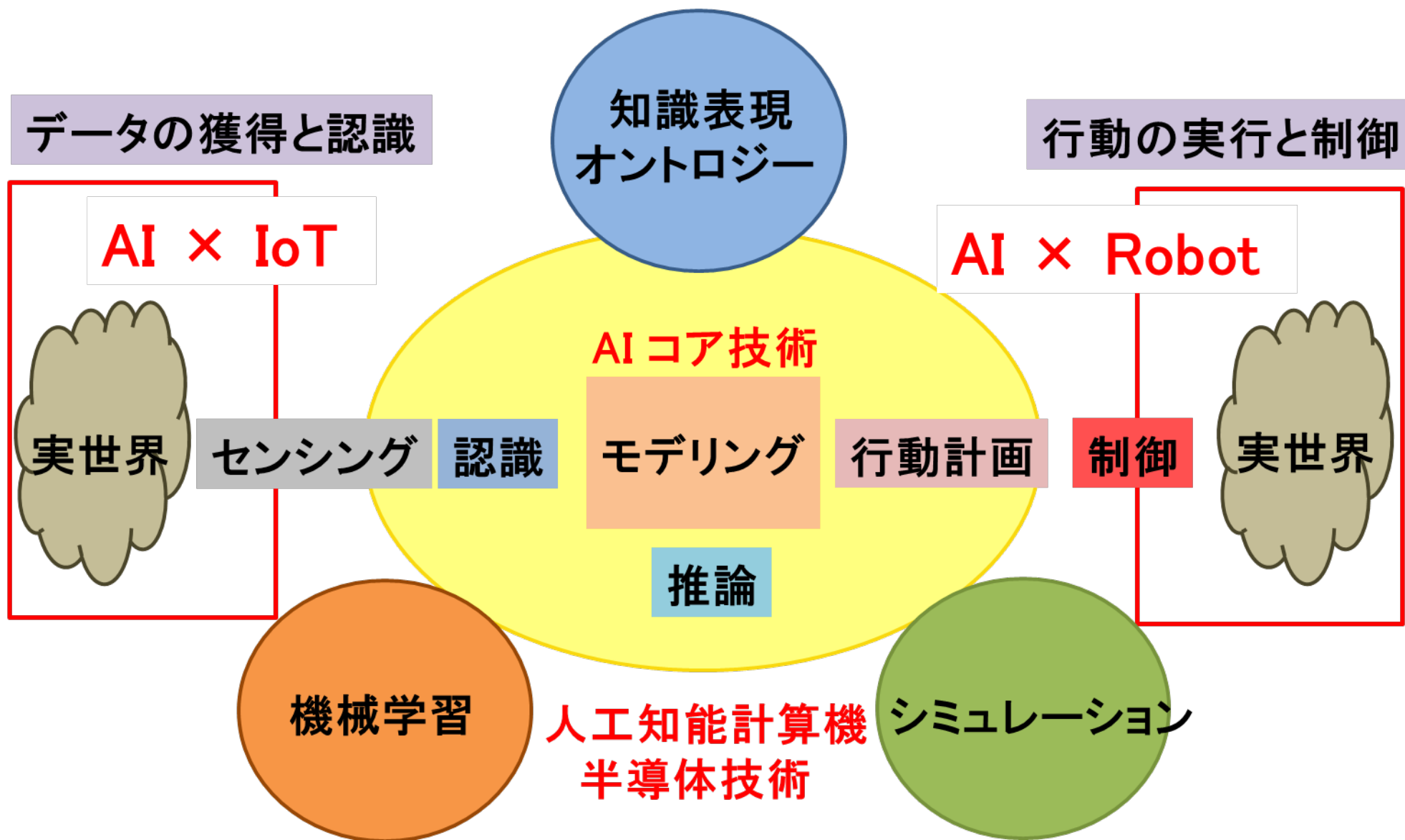


- 研究開発と社会実装の好循環の実現
 - 次世代の人工知能技術の研究開発
 - 革新的な情報処理原理の創出
 - 人工知能技術の社会実装を促進する技術の開発
 - ソフトウェア基盤
 - ハードウェア基盤
 - データ基盤
 - 研究開発拠点の形成
 - 産学官の英知を結集した研究開発拠点の形成
 - 技術の普及、社会実装のための人材育成

- 人材の集積 常勤研究員 33名 → 92名
全体では 77名 → 400名超（兼務・客員等を含む）
- 国内外の大学・研究機関との連携とグローバル化
 - クロスアポイントメント 5名、招聘研究員 18名
客員研究員 54名
 - 研究開発の再委託 13組織
 - 産総研－東工大 実社会ビッグデータ活用 OIL
 - ドイツ DFKI と MOUを締結（2017.3.19）
 - CMU、TITC とLOIを取り交わし
- 3省連携の取組み

- 民間企業との連携
 - 特定集中研究専門員 21名
 - 資金提供型共同研究 47件
 - 連携研究室・連携研究ラボ 2件
(うち 1件は 情報・人間工学領域に設置)
 - 人工知能技術コンソーシアム
(85社が参加、関西、九州支部も設立)
 - ベンチャーアウトリーチ会の開催、等
- 普及・人材育成活動
 - 人工知能セミナー・ワークショップ開催
 - NEDO/AIRC=東京大学 人工知能基礎技術／先端技術
人材育成講座の実施、等





- インターネットサービスのデータを使って
ネットサービスを改善する AI



- IoT やロボットと接続し、
実世界のデータを使い、実世界に働きかける AI

- AI のコア技術を閉じた世界で試す



- 実世界（開いた世界）のサービスの中で継続的に
収集されるデータを用いた研究開発

- AI が実世界に埋め込まれるために、人間にとって理解可能であることが重要
- 人間を代替するのではなく、人間と協働して実社会の課題を解決できる A I
- 人間と同じ記号体系（自然言語）を扱える説明可能な A I

H28年度実施項目構成 (産総研実施分)

●人材育成施策

NEDO/AIRC=東京大学
人工知能先端技術人材育成講座
NEDO/AIRC=東京大学
人工知能応用人材育成講座

③ 次世代人工知能共通基盤技術研究開発

(1) 人間行動モデリングタスク

(2) 画像解析タスク

- (a) 地理空間情報画像解析
- (b) きめの細かい動作認識

(3) テキストからの知識獲得タスク

- (a) 事故情報テキスト解析・事故予防技術
- (b) 科学文献キュレーション支援技術

(4) 対人インタラクションタスク

(5) 産業用ロボットタスク

(6) 自動運転タスク

- (a) セマンティック情報に基づく自動運転
- (b) 解釈可能な自動運転システム

効率的・横断的な技術の導入

課題への適用を通じたフィードバック

② 次世代人工知能フレームワーク研究・先進中核モジュール研究開発

(1) 次世代人工知能フレームワークの研究開発

- (a) 次世代人工知能フレームワークの研究開発
- (b) 次世代人工知能研究テストベッドの研究開発
- (c) 社会的身体性知能の共有・活用のためのクラウドプラットフォーム
- (d) ネットワーク分析と言語処理の融合による大規模文献データから技術の未来予測プラットフォームの研究開発

(2) 先進中核モジュールの研究開発

- (a) 観測・データ収集モジュール
- (b) 認識・モデル化・予測モジュール
- (c) 行動計画・制御モジュール
- (d) 自然言語テキスト理解モジュール

フレームワークによる研究推進

モジュール化

フィードバック

① 大規模目的基礎研究・先端技術研究開発

(1) 次世代脳型人工知能の研究開発

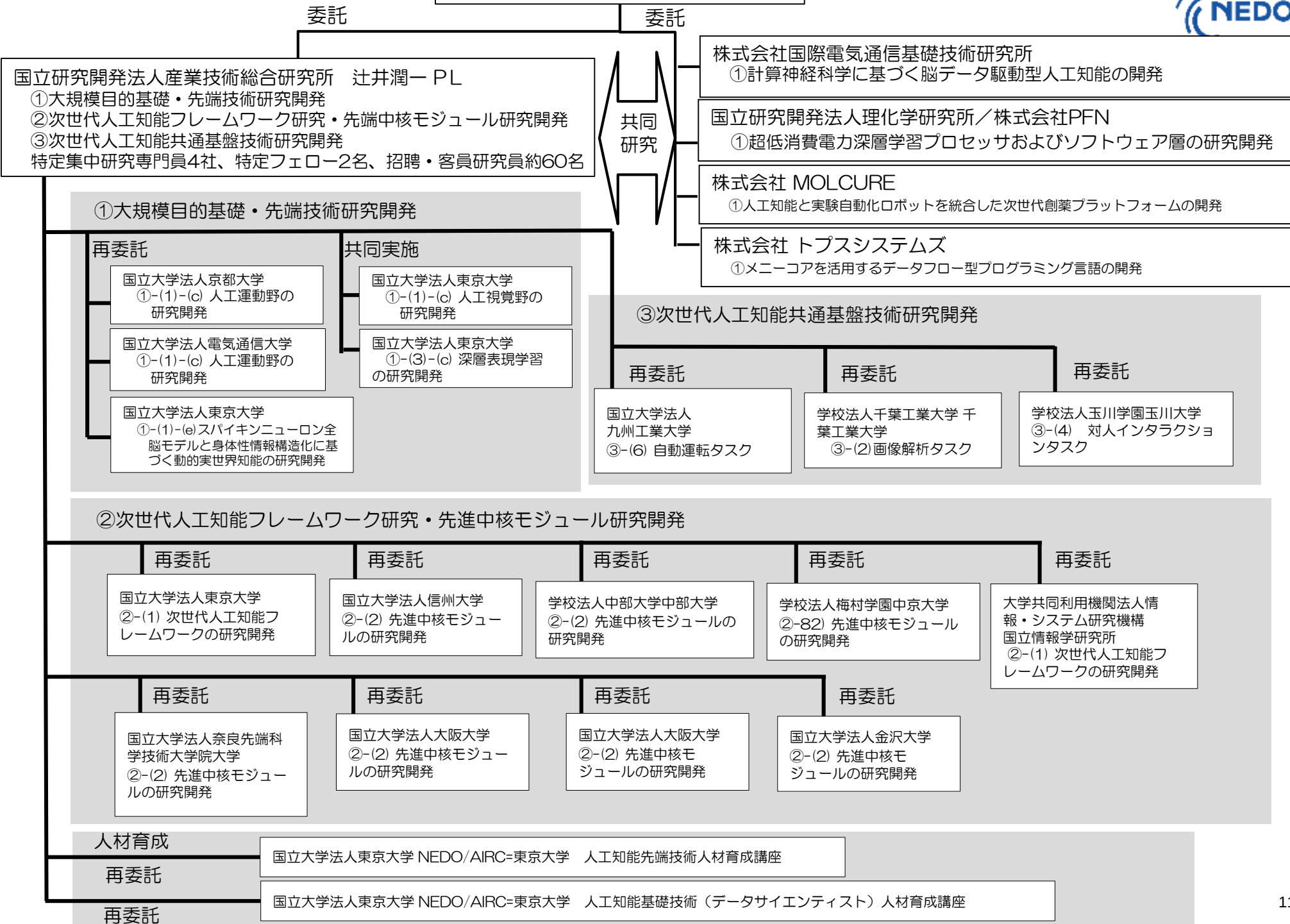
- (a) 人工大脳皮質の研究開発
- (b) 人工視覚野の研究開発
- (a) 人工運動野の研究開発
- (d) 人工言語野の研究開発
- (e) スパイキングニューロン全脳モデルと身体性情報構造化に基づく動的実世界知能の研究開発

(2) データ・知識融合型人工知能の研究開発

- (a) 自然言語理解を核としたデータ・知識融合技術の研究開発
- (b) 分散表象知識と記号的知識の相互変換技術の研究開発

(3) 機械学習および確率モデリング技術の高度化

- (a) スケーラブルな機械学習・確率モデリング手法の研究開発
- (b) 超複雑な機械学習・確率モデリング手法の研究開発
- (c) 深層表現学習技術の研究開発



①大規模目的基礎研究・先端技術研究開発 －革新的な情報処理原理の創出と高度化－

・次世代脳型人工知能の研究開発

- ・人工大脳皮質の研究開発
- ・人工視覚野の研究開発
- ・人工運動野の研究開発
- ・人工言語野の研究開発
- ・スパイキングニューロン全脳モデルと
身体性情報構造化に基づく動的実世界知能の研究開発
- ・計算神経科学に基づく脳データ駆動型人工知能の開発（ATR）

・データ・知識融合型人工知能の研究開発

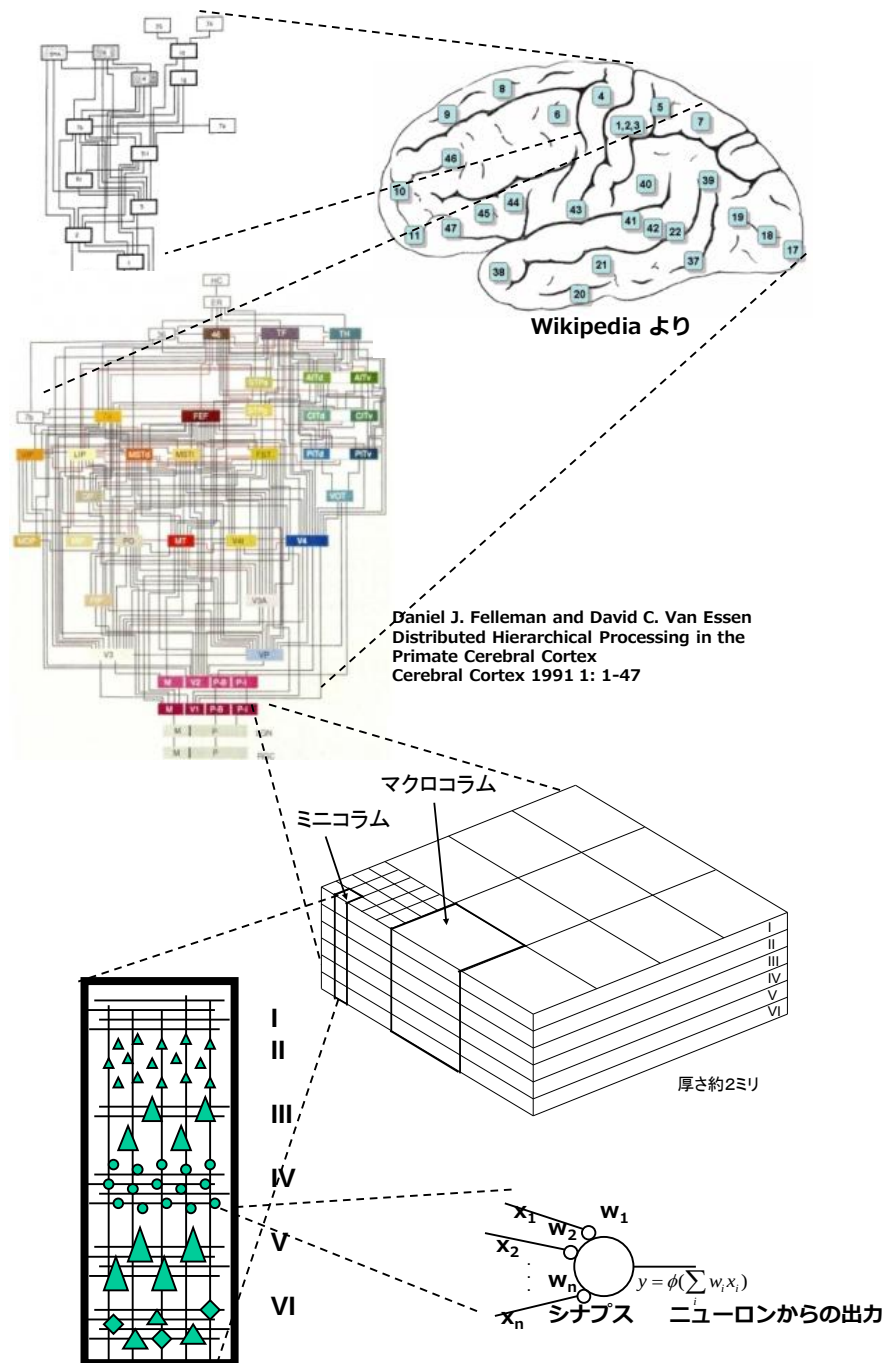
- ・自然言語理解を核としたデータ知識融合技術の研究開発
- ・分散表象知識と記号的知識の相互変換技術の研究開発

・機械学習および確率モデリング技術の高度化

- ・スケーラブルな機械学習・確率モデリング手法の研究開発
- ・超複雑な機械学習・確率モデリング手法の研究開発
- ・深層表現学習技術の研究開発
- ・超低消費電力深層学習プロセッサおよびソフトウェア層の研究開発（理研／PFN）
- ・人工知能と実験自動化ロボットを統合した次世代創薬プラットフォームの開発
（MOLCURE）
- ・メニーコアを活用するデータフロー型プログラミング言語の開発（トプスシステムズ）

人間と相互理解できる A I のための基盤的要素技術開発

- 人間の脳に学ぶ A I
 - 脳データ駆動型 A I 研究
- 人間と同じ記号体系（自然言語）を使って相互理解できる A I
 - 自然言語を通した多様なデータ・学習結果の理解
- 深層学習の高度化と記号体系を支えるイメージの獲得（記号の創発）



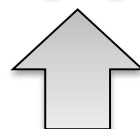
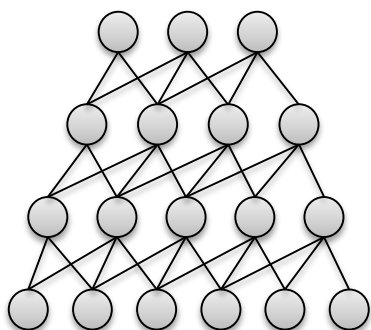
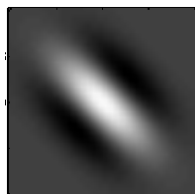
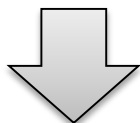
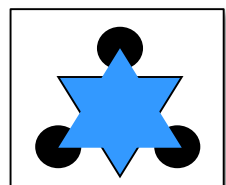
- 脳の様々な高次機能（認識、意思決定、運動制御、思考、推論、言語理解など）が、**たった50個程度**の領野のネットワークで実現されている。

**脳の情報処理原理は
まだまだ未解明な部分が多い**

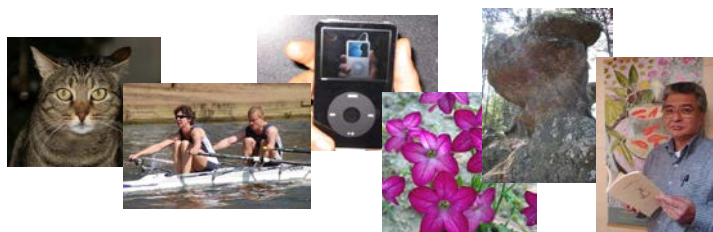
人工視覚野（ロボットの眼）

視覚認知の実現

霊長類の実験からの受容野

ヒューマノイド
上で統合類似性を
議論

ヒトに画像提示

画像特徴量による
デコーディング

ImageNetなど大規模画像DB

一般物体デコーディング
(脳活動からの一般物体認識)

CB-i Humanoid Robot

Eyes

配置
4台のビデオカメラ。

Ears

構造
2台のマイクロフォン。

Hands

手型
6自由度を有する。ウェアラブルの
パッド等を備えることが可能。
先端・中間・基端にそれぞれ
センサを有する。OFD制御。

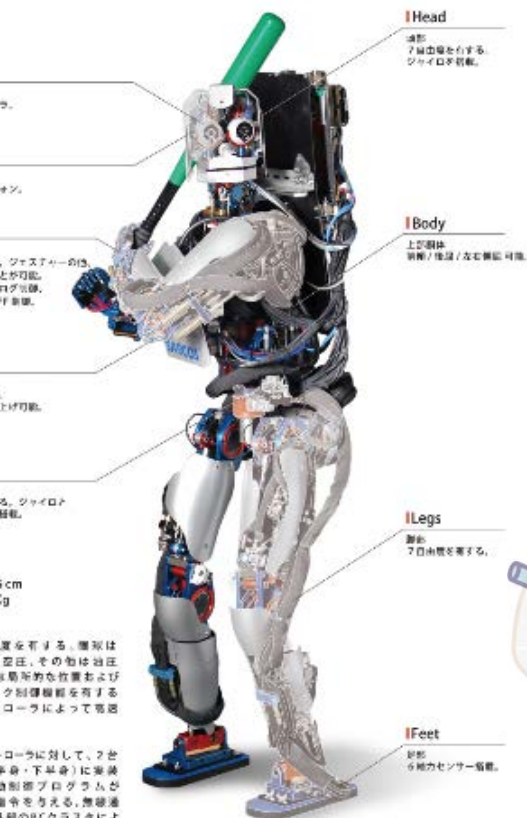
Arms

構造
7自由度を有する。
3kgの物体の持ち上げ可能。

Pelvis

構造
同様に7自由度を有する。ジャイロ
センサーを有する。Height : 155 cm
Weight : 65 kg全身51自由度を有する。関節は
電動、ハンドは空気、その他は油圧
駆動。各自由度は局所的な位置および
カフ・バック・ヒップ関節を有する
マイクロコントローラによって遠隔
に制御される。各マイクロコントローラに対して、2台
の演算PC(上半身・下半身)に接続
された上位演算制御プログラムが
実時間で運動指令を与える。無感通
信機能も有し、外観のPCクラスターに
よる制御が可能。

Humanoid Robot

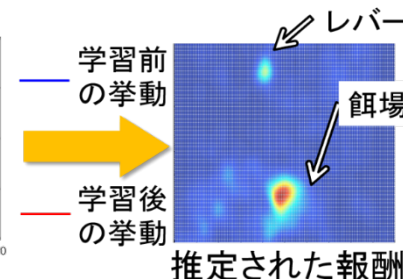
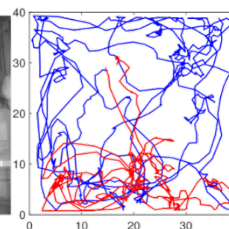
ヒューマノイド
上で統合

人工運動野（ロボットの身体）

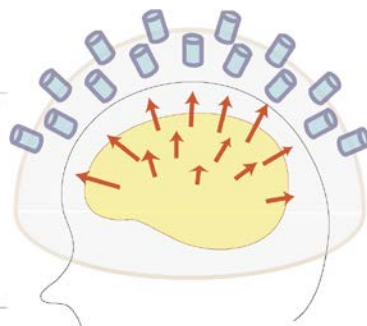
ラット



実験環境



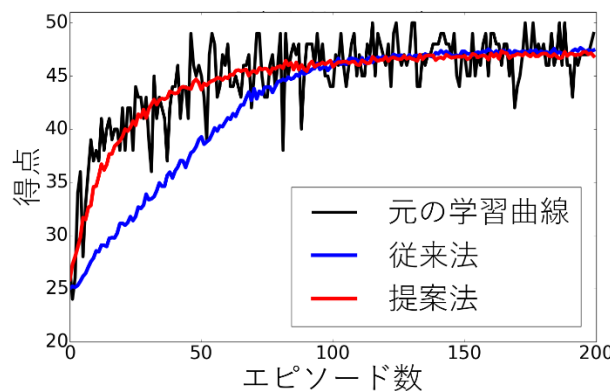
逆強化学習によるラットの報酬の推定

類似性を
議論

深層学習＋逆強化学習

模倣学習・徒弟学習

順逆強化学習＋モジュール構造

多次元特徴量による
デコーディング学習中の相手に対する
模倣(徒弟)学習

- ・ビデオの内容を表す文をAI技術(「リカレントニューラルネットワーク」)で生成
- ・従来手法ではひとコマごとに認識しているため誤認識が多く混ざること内容説明文の生成結果が不安定になったが、コマの前後の流れも考慮した「高精度被写体物体認識」を導入し、文脈の把握を安定させることで、高精度化を実現



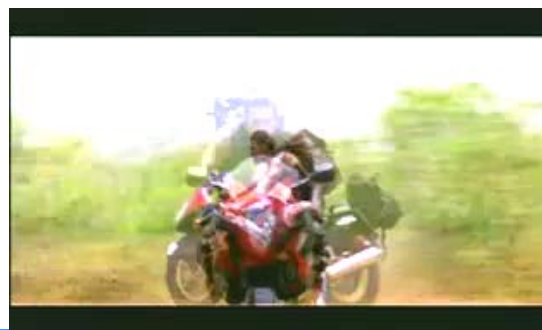
従来: A man is drinking.
 本法: A girl is doing makeup.



従来: A dog is playing with a dog.
 本法: A boy is playing with a dog.



従来: A man is riding a car.
 本法: A woman is riding a boat.



従来: A man is riding a bicycle.
 本法: A man is riding a bike.

登場人物の性別や、物体の正確な認識が可能になり、精度向上

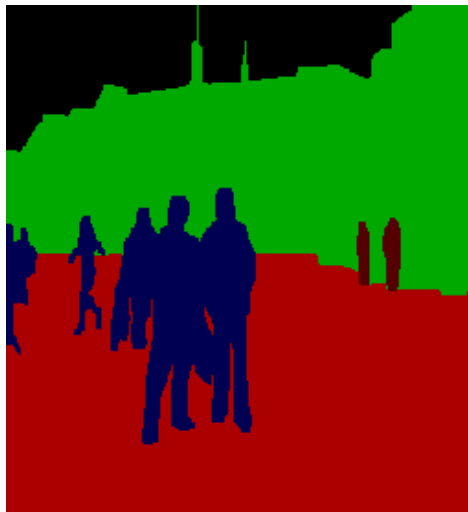
現状の先行研究では、教師データが得やすい、一般的な内容の動画が対象
 特殊・個別的な用途(工業、医療、防犯等)のビデオ認識への応用拡張

- ・画像のセマンティックセグメンテーション(認識+領域分割)の技術は、ロボット等が環境を認識して行動するために重要であるが、発展途上であり、多くのデータセットにおいて高い精度の実現が困難であった
- ・敵対的生成深層ネットワーク(Generative Adversarial Network: GAN)を画像のセマンティックセグメンテーションに適用することで、従来手法よりも高精度な認識+領域分割を実現した

敵対的生成深層ネットワークを用いることにより、識別モデルの学習のみでは実現困難な精度を実現することができる



入力画像



真値



従来手法 MIOU 0.48
(SegNet, 2015)



提案手法 MIOU 0.55

今後の課題

- ①セマンティックセグメンテーションの更なる精度向上(0.6以上を目指す)
- ②ロボット等による環境とのインタラクションによるフィードバックを導入

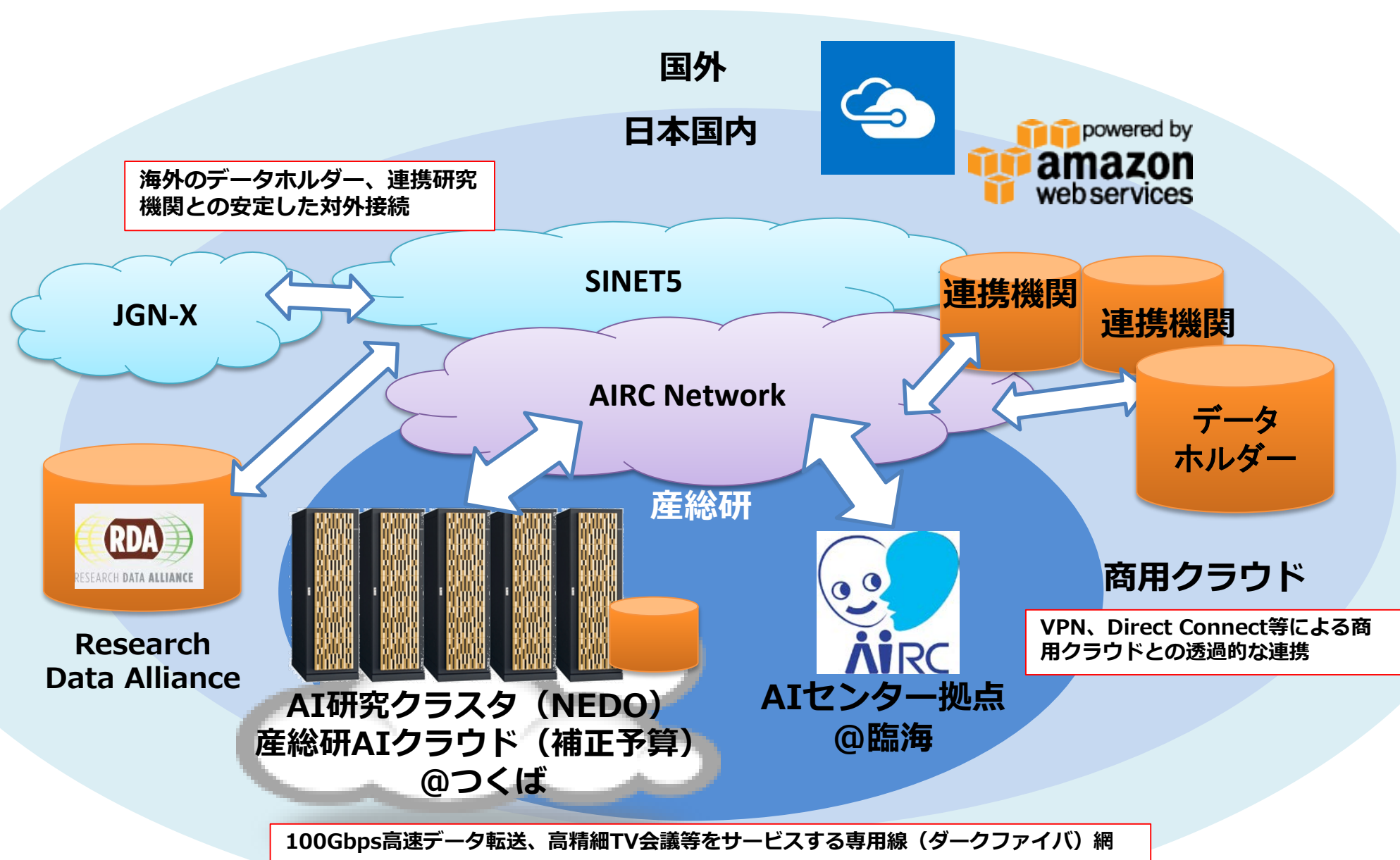
②次世代人工知能フレームワーク研究・ 先進中核モジュール研究開発 －要素技術の利活用の促進－

・次世代人工知能フレームワーク研究

- ・次世代人工知能フレームワークの研究開発
- ・次世代人工知能テストベッドの研究開発
- ・社会的身体性知能の共有・活用のためのクラウドプラットフォーム
- ・ネットワーク分析と言語処理の融合による大規模文献データから技術の未来予測プラットフォームの研究開発

・先進中核モジュールの研究開発

- ・観測・データ収集モジュール
- ・認識・モデル化・予測モジュール
- ・行動計画・制御モジュール
- ・自然言語テキスト理解モジュール



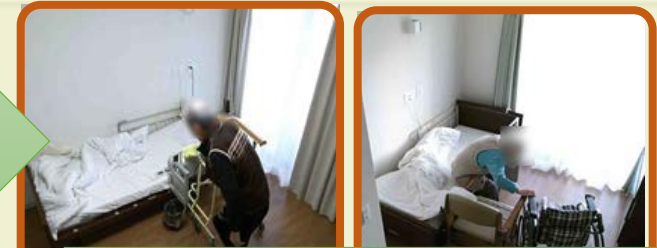
- ・実際の現場を想定した**仮想実験のための産総研内リビングラボ**と、現場での**実証フェーズ**を支援する**サテライトリビングラボ（10か所）**環境を構築
- ・基礎開発から実証研究までをシームレスにつなぐ国際的にも類のないリビングラボ
- ・収集したデータに基づき、年齢軸だけでなく生活機能軸で生活現象を整理

実世界・実問題と繋がったスマートリビングラボ

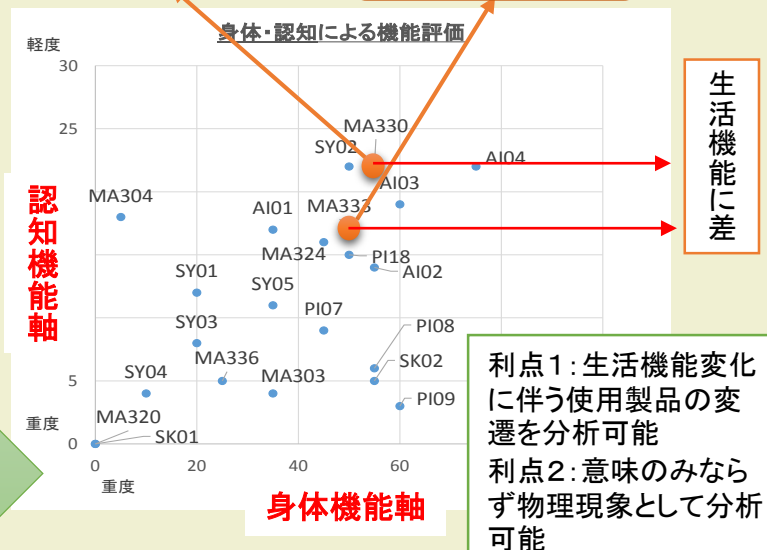


意味データ
(例. アンケート、
行動記録)

センサデータ
(例. RGBD (カメラ)、
荷重センサー)



行動ライブラリ(リビングラボ応用例)



今後の課題

- ①年齢軸から生活機能軸へのパラダイムシフト（製品・サービス設計）のためのオープンデータベース構築。
- ②AIによるプリシジョン・ケア（高精度ケア）への展開：個人の生活機能変化に適応した見守り技術、リスクマネジメント技術の開発

- ・クラウド型の仮想現実（VR）環境を構築し、クラウドソーシングによる大規模長時間対話実験を可能に
- ・数百人が同時に参加できる環境で、1万時間オーダーの対話経験をクラウド上で共有
- ・ロボットの社会的知能開発の飛躍的効率化

UnityによるVRユーザインタフェースと、ROSによるロボットプログラムの開発環境を統合することで従来までのソフトウェアを改良することなく、実機ロボットでの対話実験をVR上で再現可能

【従来の手法】



ハードウェアトラブル多
被験者集めのコスト大

対話経験の共有：困難
長くても数時間ほどの実験が限界

【新しい手法】



ログイン

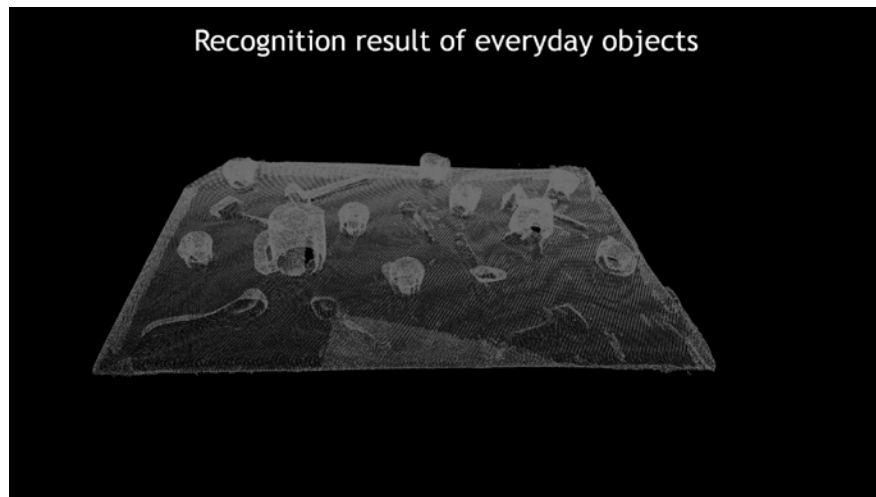
ハードウェアトラブル無し
被験者は時間場所を選ばず参加

1万時間オーダーの対話経験を
クラウド上で共有可能

今後の課題

- ①対話ロボットから生活現象モデリングへの展開（介護/保育活動支援等）
- ②コンペティション形式の社会実験による定量的な効果の検証

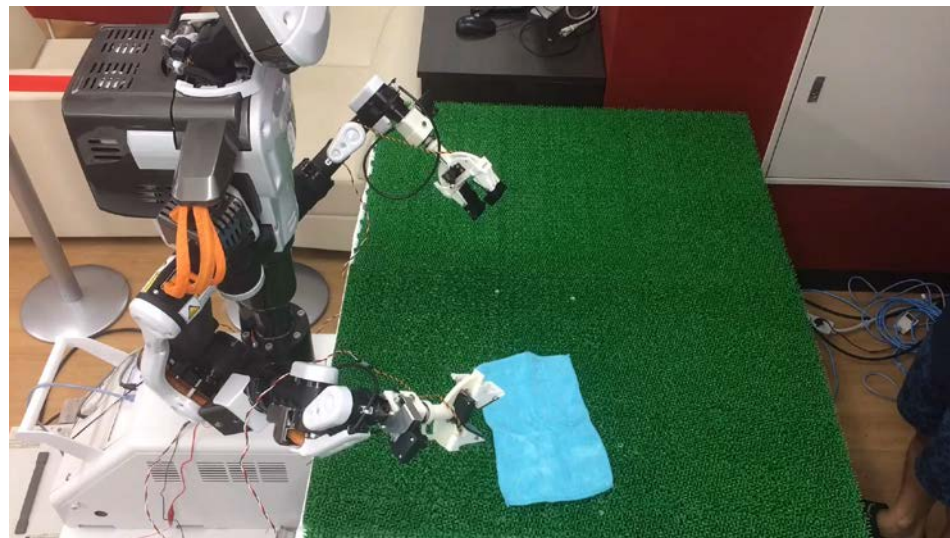
3Dデータからの日用品のカテゴリと機能の同時認識



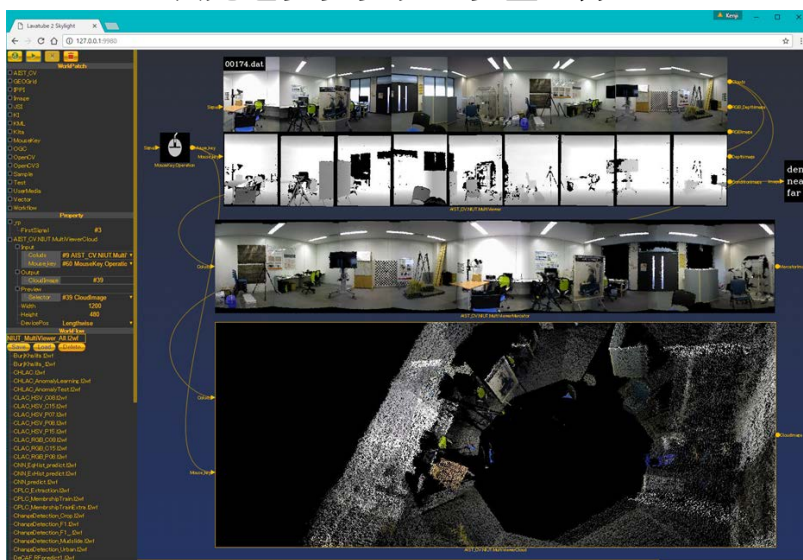
物体認識成功率：92%

第21回知能メカトロニクスワークショップ(IMEC2016) 優秀講演賞

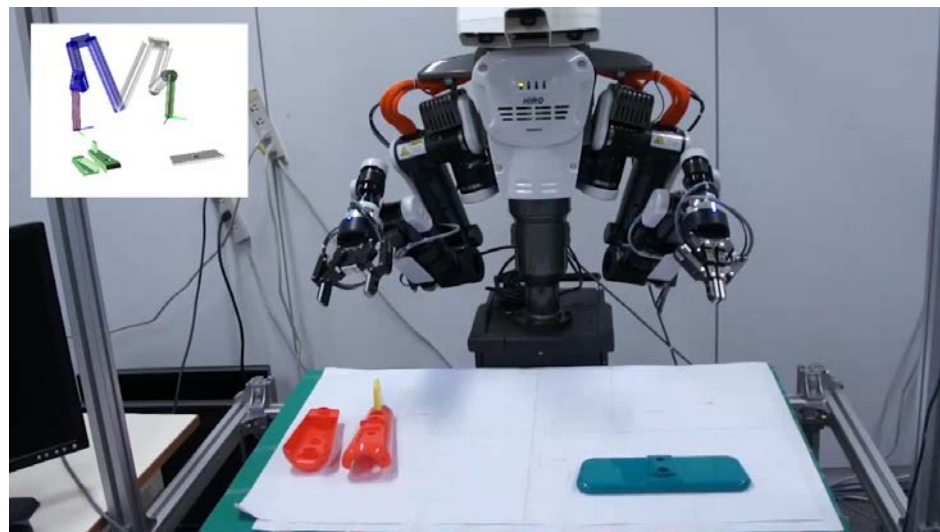
柔軟物操作の模倣学習



3次元センシングモジュール



組立作業動作の自動計画・生成



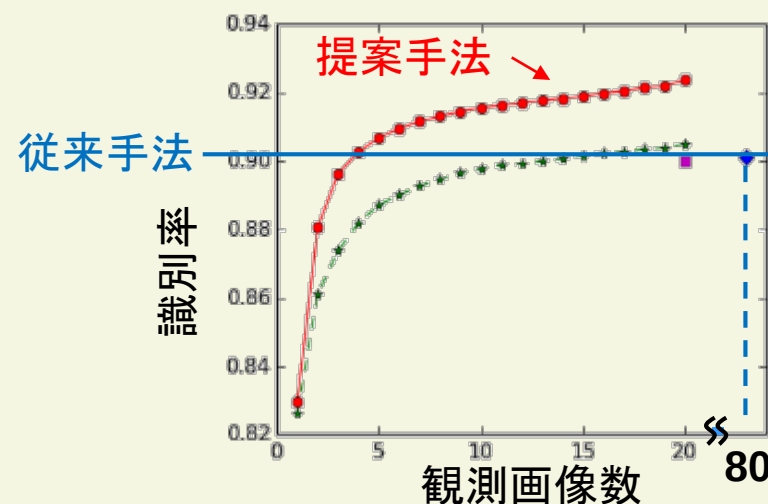
ロボットプラットフォーム Deep Learning Robot



三次元物体検索の国際的コンペ (SHREC2017) において、**本開発アルゴリズムが第1位を獲得**

少量の観測画像数で従来手法の精度を上回る新しい物体認識技術を開発

90%⇒92%、80枚⇒20枚以下



★12種類、132個、21,120枚からなる日用品画像データセットを構築し、学習



－人工知能技術・システムの評価・検証－

- 標準タスクの設定
 - 人間行動モデリングタスク
 - 画像解析タスク
 - テキストからの知識抽出タスク
 - 対人インタラクションタスク
 - 産業用ロボットタスク
 - 自動運転タスク
- タスクの性能評価手法検討のためのベンチマークデータセット構築
- 標準化に向けた検討

サービス現場での人間行動モデリング

AI応用システム
実証フィールド



技術移転と共同研究での利用実績



現場オペレーション

AIが人を理解

人間行動モデリング
行動予測
(標準タスク)

標準的機能の提供
次世代人工知能API
(POSEIDON)

アクティブ
ビッグデータ収集
(標準データ生成・管理)

分散データ
統合システム
(MCDataBinder)

人間行動
モデル

AI応用システムとして
社会実装

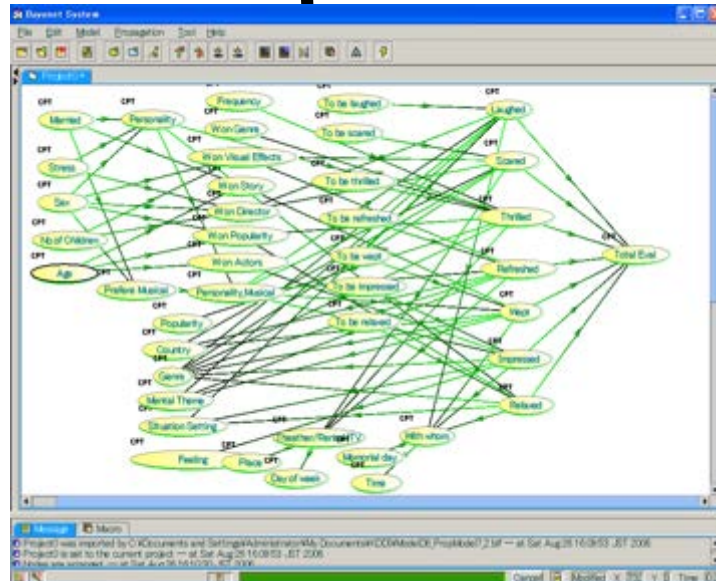
相互理解を
深める

マネジメント

データ・知識
循環駆動

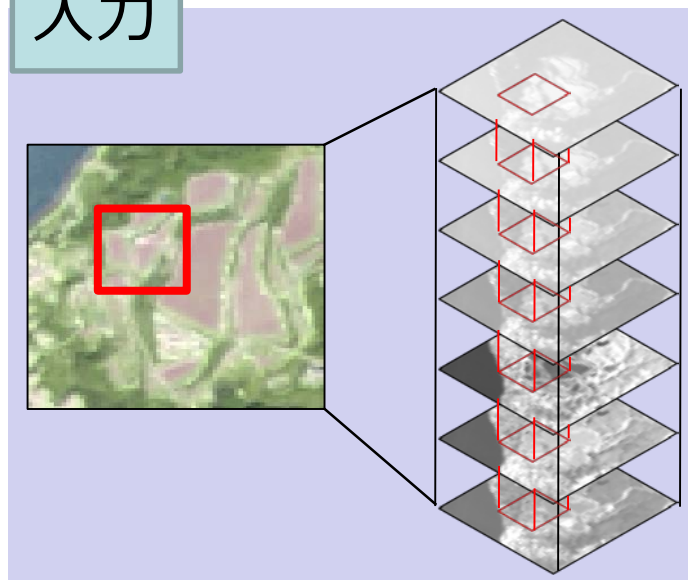
データ・知識融合
機械学習

人がAI(計算過程)
を理解



先進中核モジュール開発
(PLASMA-N1: PLSA+BN)

入力



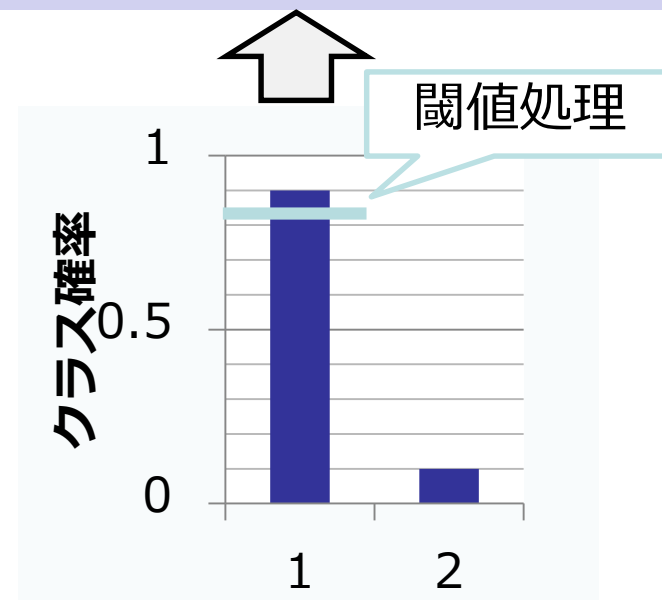
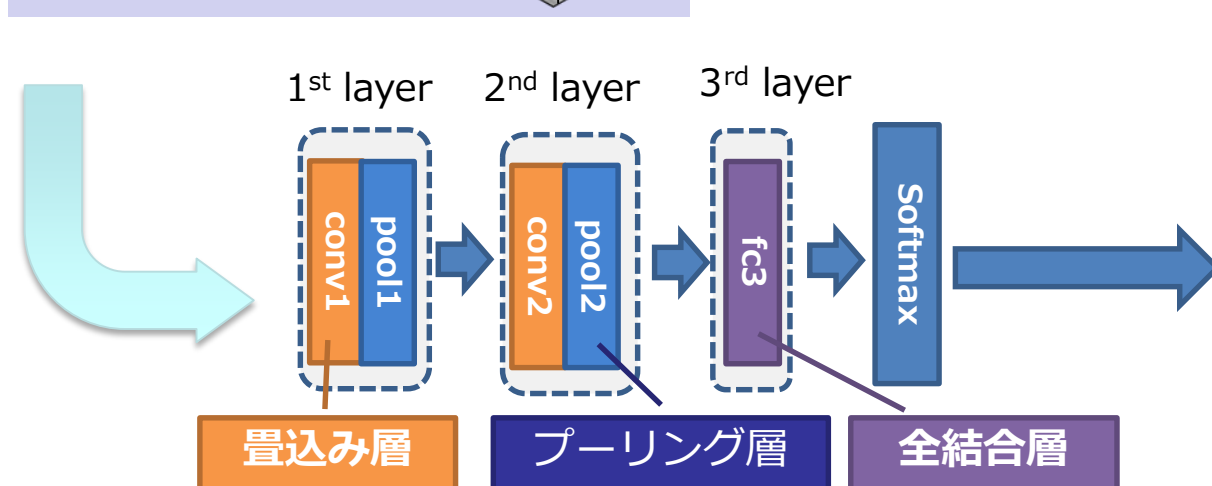
観測波長領域が異なる衛星画像を16x16のセルに分割

約5000枚の教師セルを人間が入力し学習

出力



メガソーラーの検出結果



ヒトの多様な動作を高精度で認識する技術の研究開発

- 日常的なシーンにおけるヒトの動作（座る、部屋から出る、ご飯を食べる、水を飲む、抱き合う、など）100種類を定義
- この100種類の動作 x 動画1000本 合計100,000本の動画データセットを構築（中）
- このデータセットと深層学習で精度の高い認識エンジンを開発

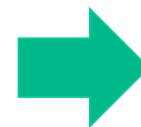
日常生活シーンを理解・支援できる人工知能へ（介護、保育、家事手伝い） *AI for Human Life*

深層学習による認識技術開発の展望： 大規模データセットによる訓練が高性能を引き出す

動画データセットを構築

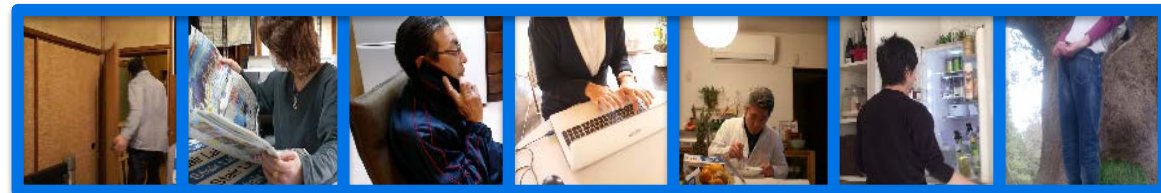
【従来の動作データセット】

動作の種類が少ない or
動画の数が少ない



【構築中の動作データセット】

日常動作の種類は約2倍
種類別の動画数は約7倍



動画認識エンジンを開発

CNN+LSTM型の認識エンジンを開発

開発中の認識エンジンの精度

今後の課題

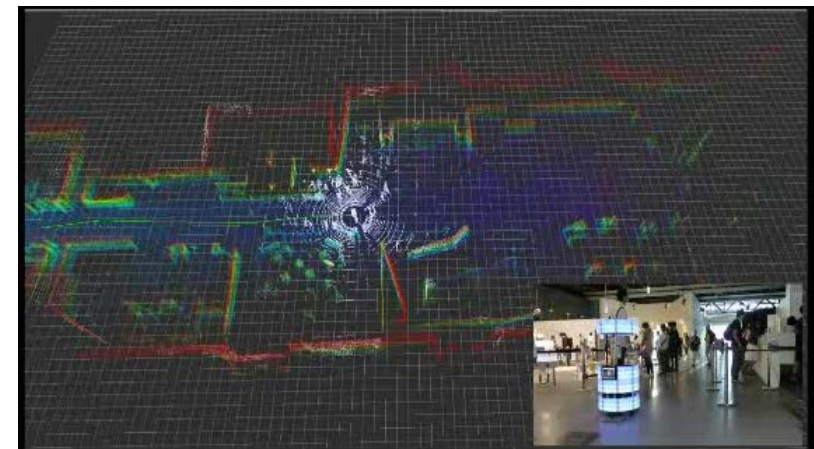
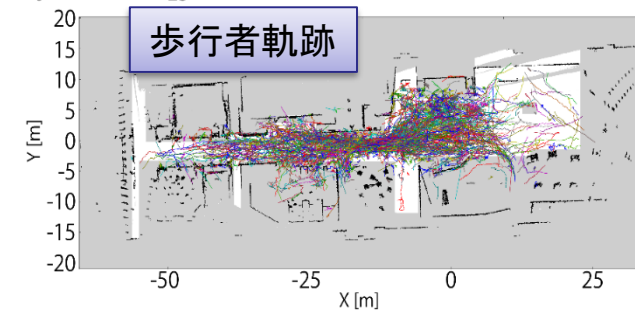
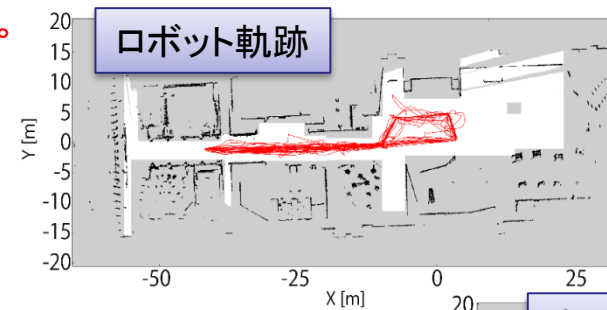
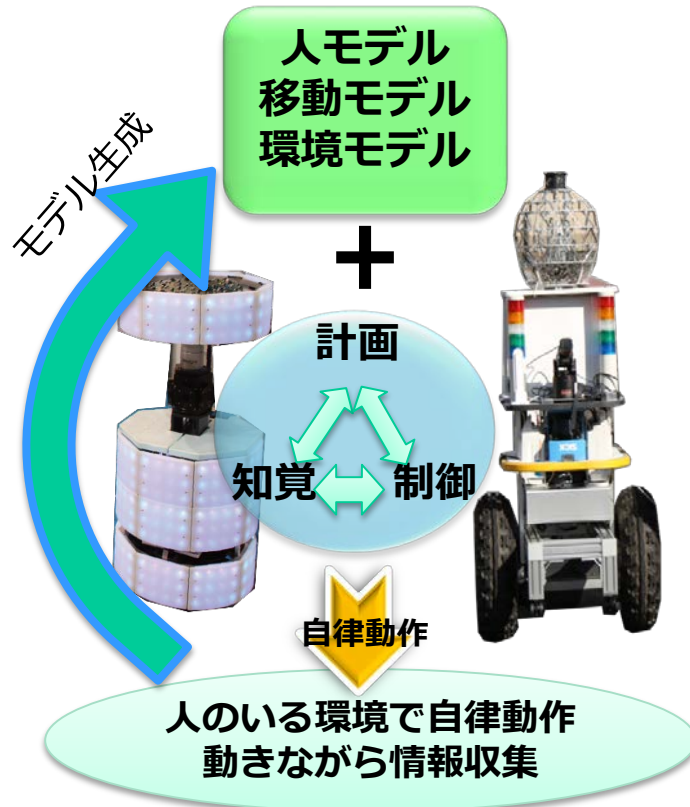
- ① 認識精度の更なる向上
- ② 動作シーンを文章表現する手法の開発

構築中のデータセット	45.6%
UCF101	86.04%
HMDB	52.87%

従来のデータセットよりも難易度高

- ・ ロボットが自走しながらセンシング（LIDAR等による）を行い、周囲の物体・人を認識し、時間経過で変化する意味付けされた 3 次元地図を作製。

- 環境センシング + 自己位置推定
⇒ **ダイナミック・セマンティックマップ**



人の移動の観測とモデル化から「通路」の概念をその場で学習

- ・ 通常は、形状地図（部屋の形状+障害物位置）を与える必要がある
- ・ 人が通っている場所 = 通れる場所と学習

1) 論文発表、学会招待講演、依頼講演等

- ・ ICPR, EMNLP, COLING, IEEE BigData, ICONIP 等の著名国際会議での発表、招待講演 等

2) ソフトウェアモジュール・プラットフォームの構築

- ・ リビングラボ、解釈可能な人間行動モデリングモジュール、次世代自販機
- ・ 一般物体認識エンジン、動画認識モジュール、衛星画像からの地物認識モジュール
- ・ 自然言語処理モジュール、等を構築中

3) 学習用・ベンチマーク用データベースの構築

- ・ 日用品 3D データベース、日常生活動作動画データベース、
- ・ 衛星画像データベース、等を構築中

4) 新聞報道・受賞等

- ・ 産経新聞 2016/10/31朝刊2面 自動運転技術、等
- ・ 情報処理学会インタラクシオン2016 インタラクティブ発表賞、第21回知能メカトロニクスワークショップ (IMEC2016) 優秀講演賞、等

5) 研究インフラ整備

- ・ AI研究用クラスタ Nairobi の構築・運用
- ・ 産総研内リビングラボの構築・運用

- 数多くの研究開発テーマの中から、主な成果について
以下のようにグルーピングしてご紹介
 - 大規模目的基礎研究
 - AI 計算基盤
 - 招待講演「Brain Inspired Computing」
 - AI for 科学技術研究
 - AI for ロボット
 - AI for 生活支援
 - AI for 地理空間情報