

ディープラーニング（深層学習）を中心とする 第 3 世代型人工知能の限界を探る調査研究

報告書

平成 3 0 年 5 月

目次

目次	i
1. 調査研究の概要	5
1.1 実施経過	5
1.2 業務実績の説明	5
2. 実態分析	9
2.1. 対象データと分析の方法	9
2.1.1. 評価分析作業について	9
2.1.2. 事例評価モデルの詳細について	9
(1) 知識カテゴリ	9
(2) 発展段階レベル	10
(3) 研究開発ステージ	11
2.1.3. 評価分析作業の流れ	11
2.1.4. 事例データの概要	12
(1) 日本経済新聞	12
(2) 人工知能学会誌	13
(3) Preferred Networks 社のホームページに開示された事例	13
(4) 政府の取り組み事例 (NEDO)	13
2.2. 日本経済新聞記事	14
2.3. 人工知能学会誌	16
2.4. Preferred Networks 社の事例	17
2.5. 政府の取り組み (NEDO 事例)	17
2.6. 政府の取り組み (全般)	18
2.6.1. 調査分析の方法	18
2.6.2. 「人工知能技術戦略会議」	19
2.6.3. 「日本再興戦略」の概要	24
2.6.4. 第5期科学技術基本計画 (平成28年1月22日閣議決定)	25
2.6.5. 科学技術イノベーション総合戦略2017 (平成29年6月2日 閣議決定)	26
2.6.6. 新産業構造ビジョン (平成29年5月30日 産業構造審議会新産業部会事務局)	27
2.6.7. 「人工知能と人間社会に関する懇談会」報告書 (170324 内閣府←160530)	28
2.6.8. まとめ	29
3. 得られた知見のまとめ	31
3.1. 評価分析結果のサマリについて	31
3.1.1. 全体サマリ	31
3.1.2. 事例パターン	32
3.1.3. 各軸間のクロス集計	34
(1) 「発展段階レベル」と「知識カテゴリ」	35

(2) 「発展段階レベル」と「研究開発ステージ」	36
(3) 「知識カテゴリ」と「研究開発ステージ」	38
3.1.4. その他	39
(1) 発展段階レベル	40
(2) 知識カテゴリ	40
(3) 研究開発ステージ	40
(4) 記事の知識カテゴリ別	41
3.2. 現状取り扱われている人工知能の実態について	42
3.2.1. 自然システム・人工的物理システムの状況（ハード1）	43
3.2.2. 自然システム・人工的物理システムの状況（ハード2）	43
3.2.3. 人工的抽象システムの状況（ソフト1）	44
3.2.4. 人工的抽象システムの状況（ソフト2）	45
3.2.5. 人間活動システムの状況（ソフト3）	47
4. レベル4人工知能の限界	51
4.1 制御性の問題	52
4.1.1 スピードのロジックが拡散させるブラックボックス	52
4.1.2 「蓋然性による正確性」というチャレンジ	53
4.1.3 「想像をはるかに超える長い道のり」	57
4.2 透明性の問題	59
4.2.1 人工知能に置換されるトレーダー	59
4.2.2 人工知能に置換されないファンドマネージャー	60
4.2.3 「責任者はどちらにいらっしゃいますか？」	63
4.2.4 「疑似ソフト化されたハード」：トレーディング・プログラム	67
4.2.5 民主主義の人工知能、権威主義の人工知能	69
4.3 創発性の問題	72
4.3.1 ピークは過ぎたFinTech、次のブームを狙う？「HR テック」	72
4.3.2 「アルゴリズムック・バイアス」－偏見を内蔵する人工知能	74
4.3.3 得意分野は時代遅れの恐怖モジュール	77
4.3.4 不得意分野はダイナミック・ヒューマニティ	80
5 次世代人工知能を構想するためのアプローチ	87
5.1 哲学的アプローチ－認識論と知識論の精査	87
5.2 事例分析によるアプローチ	87
5.3 脳科学や心理学からのアプローチ	90

図一覧

図 1 対象の知識カテゴリに登場する各項目の関係	9
図 2 評価分析作業の流れ	11
図 3 登場する語句の共起状況	16

図 4 取り上げられた事例のパターン（人工知能のレベル別）	42
-------------------------------	----

表一覧

表 1 対象の知識カテゴリの説明	10
表 2 人工知能の発展段階レベルの説明	10
表 3 研究開発ステージの説明	11
表 4 本件評価分析作業の内容	12
表 5 新聞記事 事例の概要	12
表 6 対象となる人工知能学会誌の概要	13
表 7 Preferred Networks 社事例の概要	13
表 8 NEDO 調査研究採択テーマにおける事例の概要	14
表 9 日本経済新聞に掲載された人工知能関連記事の分類	14
表 10 特徴評価を実施した人工知能学会誌の事例の一覧	16
表 11 収集事例の分類	17
表 12 評価を実施した NEDO 採択案件の分類と抽出事例数	18
表 14 新聞記事における各カテゴリ毎の 3 軸内の登場割合	40

1. 調査研究の概要

1.1 実施経過

(1) 調査実施者

非常勤上席研究員： 平澤 洽

常勤主任研究員： 笠井 祥、 林 隆臣

常勤研究員： 山本智史

非常勤主席研究員： 宮下 永、 宮林正恭、 大川晋司、 太田与洋、 関 直康

田中好雄、 中崎孝一、 藤江健介、 藤本昭雄、 伊藤真理

山田典子、 依田 孝 力 利則

(2) 調査実施経過

本調査研究は、人工知能分野に興味を持つ17名という大人数で一部研修を兼ねて実施した。提案書は平成29年12月11日から平成30年2月26日まで5回にわたる全体会議で検討した。実質的調査分析は3月12日から分担して開始し、途中2回の進捗状況の報告会を持ち、4月23日にまとめの議論を経て分担部分の執筆に入った。以後は、日本経済新聞、人工知能学会誌、Preferred Networks社、政府の取り組みに関し6月18日まで2回の検討機会を経てドラフト原稿を作成した。その後は、平澤、笠井、林、山本、中崎が担当を分け、個別に検討を深めた。

1.2 業務実績の説明

(1) 調査の目的

人工知能の第3次ブームを支える深層学習手法による取組みを整理し、成果を生み出す原理について追究することにより、深層学習手法の限界を明確にする。

(2) 調査の概要

人智を凌駕するかにみえる人工知能の多様な活躍事例を取り上げ、その成果を知識論の原理にさかのぼって整理し、第3世代型人工知能の得意分野と不得意な領域とを見分け、第3次ブームを先導する深層学習手法の可能性についてまとめた。

(3) 調査研究の内容

1) 対象事例

- ・2017年日本経済新聞記事からAI等のキーワードを含む804件
- ・2017年人工知能学会誌所載論文121件
- ・Preferred Networks社のホームページに開示された実施事例71件
- ・政府の取り組み事例であるNEDO採択テーマ57件

2) 分析の方法

上記対象事例を次の3軸の各評定区分により特徴付ける。

- ・人工知能を作動させている対象の知識のカテゴリ（「知識カテゴリ」）
- ・使用する人工知能手法の発展段階レベル（「発展段階レベル」）
- ・人工知能を適用する研究開発ステージ（「研究開発ステージ」）

3) 事例の属性の特徴

事例を 3 軸の属性によって区分し事例特性の分布状態を数量的に把握する。以下のような特徴が見られる。

- ・ディープラーニングを適用している事例は各データソースともに多いが、NEDO 案件では人材養成事例を除きすべてがディープラーニングであり、また PFN 社では機械学習の適用も見られるがディープラーニングが圧倒的に多い。これに対して、人工知能学会誌では機械学習とディープラーニングが拮抗しているがディープラーニングがやや多く、日本経済新聞では逆に機械学習のほうがやや多い。

- ・ディープラーニングを適用した研究開発ステージを見ると、データソースによって際立った特徴が見られる。人工知能学会誌では基礎研究ステージであるのに対して PFN 社では市場を意識したステージに偏在している。NEDO では応用研究ステージである。これに対して日経の記事では製品開発のステージにピークがあり山形に分布している。

- ・知識カテゴリに関しても興味深い分布が見られる。日経の事例は人工抽象システムが多く、PFN と NEDO は自然・人工物理システムに偏在し、また学会誌では人工抽象システムと人間活動システムに集中している。

- ・PFN、NEDO、学会誌にはディープラーニングの人間活動システムへの適用は見られない。

- ・日経の記事は内容によって 22 のカテゴリに分類され、類似内容の記事の特性を比較した。件数では業務支援が最も多く、音声認識、フィンテックがそれに続く。件数は多くないが物流や自動工場にも特色が見られる。

4) ディープラーニングの限界

ディープラーニングの適用事例を巡って、関連文献を検索し検討を深めた。

- ・共通して指摘されていることは、ディープラーニングによって抽出される特徴量が、注目している特性との間に何らかの相関関係が有るが、どのようはメカニズムでそれが発現されるのかについては不明のままである。つまり因果関係を解き明かすことが出来ない。

- ・人工物理システム（物流・自動走行）に対しては、ディープラーニングでは本質的な解決は見通せないものの、補完的な方法により漸進的に実装可能な機能を拡大している。

- ・人工抽象システム（囲碁・金融）に関しては、囲碁では成功したものの、金融では課題解決が見通せていない。

- ・人間活動システム（人事）への適用例は多いが、課題解決は見通せていない。

- ・知識論が示す原理によれば、以下の説明が可能である。対象知識に普遍的な内在原理が存在する自然システムや人工物理システムに属するデータのみからなる系では、ディープラーニングにより抽出する特徴量には普遍性が備わっているが、人間活動システムでは普遍的内在原理の存在は期待できず、人間活動システムに属するデータ集合から抽出される特徴量には普遍性は期待できない。また人工抽象システムの場合は特定のモデルに属するデータのみからなる系では特徴量に期待できるが、異なるモデルに属するデータが混在している場合には特徴量は普遍性を持ち得ないと考えられる。

5) 次世代人工知能を構想するためのアプローチ

第 3 世代型人工知能のホープと目されているディープラーニングの限界が見えてきた現在、数理的な機械学習とは異なる発想の下で、次世代の人工知能を構想するヒントが得られる。

- ・異なる哲学の下での認識論・知識論によるアプローチ。科学主義の哲学の下では人間活動システムに適用可能な認識論・知識論を構想することが困難だとすれば、現象論等の異なる哲学の下で新たな認識論・知識論の構想を試みる。
- ・AIによる把握が科学主義の哲学の下では困難な事例について検討を深める。人文学の領域やVR仮想現実やAR拡張現実等を含む先進的な映像の領域等が考えられる。
- ・脳科学や心理学等の生物的な実体を担っている領域とAIとの関係に切り込むことも考えられる。

2. 実態分析

2.1. 対象データと分析の方法

現状の深層学習に代表される人工知能に関する様々な取り組みについて、その成果を生み出す原理について追及するために、以下にあげる評価分析モデルを用意し、収集事例の評価と分析を実施した。

対象となる事例は局面の異なる次の4種とした。

- ① 日本経済新聞
- ② 人工知能学会誌
- ③ Preferred Networks 社のホームページに開示された事例
- ④ 政府の取り組み事例 (NEDO)

※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) の「次世代人工知能・ロボット中核技術開発／次世代人工知能技分野」(調査研究) 採択テーマ

2.1.1. 評価分析作業について

入手した各事例の特徴を把握するために属性を評定した。「人工知能を作動させている対象の知識のカテゴリ」、「使用する人工知能手法の発展段階レベル」、「人工知能を適用する研究開発ステージ」の3軸を用い、作業メンバーが分担して実施した。3軸を端的に表現すると以下ようになる。

- ① 知識カテゴリ
- ② 発展段階レベル
- ③ 研究開発ステージ

2.1.2. 事例評価モデルの詳細について

3軸の具体的な内容について以下に説明する。

(1) 知識カテゴリ

人工知能を作動させている対象の知識カテゴリー					
実体的概念による対象区分					
チェックリストの類概念による区分		自然工学システム	境界領域	人間・社会システム	
	物理システム 人工的システム と 自然システム	ハード1		ハード2	ハードな対象
	抽象システム 人工的システム	ソフト1 (第1類)		ソフト2 (第2類)	ソフトな対象
	人間活動 システム	(第4類)		ソフト3 (第3類)	

図 1 対象の知識のカテゴリに登場する各項目の関係

知識カテゴリ軸としては、知識論に基づく知識のカテゴリ区分を用いる。チェックランドの知識の類概念を用いると、知識は図1のように区分できる¹。横軸は対象の実態としての概念で、意思を持った人間が関与するかどうかで区分する。また、縦軸はチェックランドの類概念による区分である²。図1の各部分の特色を表1にまとめる。

表 1 対象の知識カテゴリの説明

知識カテゴリ	対象	特徴
ハード1	自然や機械のような装置	対象に内在する法則がある
ハード2	意思や価値観を持った人間を含む	内在する法則は無い。対象により異なり、時間とともに移ろい行く
ソフト1	自然や機械を模した外在的モデル。外部から眺めたモデル	シミュレーションの結果を照合する実体に法則性がある。
ソフト2	人の行動や組織・社会を模した外在的モデル。外部から眺めたモデル	シミュレーションの結果を照合する実体が不均一で不定。しかしその程度が低い対象を仮定的な法則性を立てて扱う
ソフト3	意思や価値観という内面を含むモデル	科学主義の哲学の立場から確定的に把握することが困難。
その他	上記の知識カテゴリに分類できないもの（たとえば「超越的概念」）	

（2）発展段階レベル

事例に登場する人工知能の発展段階レベルを人工知能が制御する対象とその仕組みに基づき表2の4つに分類する。分類に当たっては、制御する対象や仕組みを示す制御系のキーワードを抽出し、その判断に用いる。

表 2 人工知能の発展段階レベルの説明³

	特徴
レベル1	単純な制御系 機械仕掛けやフィードバック機構を備えている 限定的な選択肢しか用意されていない
レベル2	推論・探索機構や知識ベースを内蔵し、与えられたルールや枠組みの下で、多様な対応が可能
レベル3	教師付きデータ（データの意味が与えられている）による機械学習機能を備えている
レベル4	ディープラーニングを活用 教師無しデータから特徴量を抽出し、外部からの指示なしに対応する機能を備えている
その他	上記のレベルに分類できないもの（たとえば複合的な手法等）

¹ 「ソフト系科学技術の研究開発の現状及び今後の展開方向についての調査」 p.38、（財）政策科学研究所（1988）

² Peter Checkland（高原康彦、他 訳）「新しいシステムアプローチ」 p125,137 オーム社（1985）

³ 松尾 豊「人工知能は人間を超えるか ディープラーニングの先にあるもの」より

(3) 研究開発ステージ

前述の2軸により評価された事例が、人工知能により何（ないしはどのような機能）を直接的に実現しようとしている研究開発ステージについて、次の表にあげる6つに分類する。

表 3 研究開発ステージの説明

ステージ	ステージ区分の特色	略称
科学研究	識（知）るための知識	ステージ1
技術研究・技術開発	実現する（ないし作る）ための知識	ステージ2
開発研究	機能を発揮することを確認するための試作	ステージ3
製品開発	販売することを目的とした製品開発	ステージ4
競争力開発	市場競争を有利にするための改善	ステージ5
基盤研究・開発	共通基盤を担う（ないし支持する）知識や製品	ステージ6

2.1.3. 評価分析作業の流れ

評価分析作業は以下の流れにより実施した。なお、作業に当たる前に、作業メンバーの人工知能に対する前提知識の共有を図るための事前ゼミを実施している。

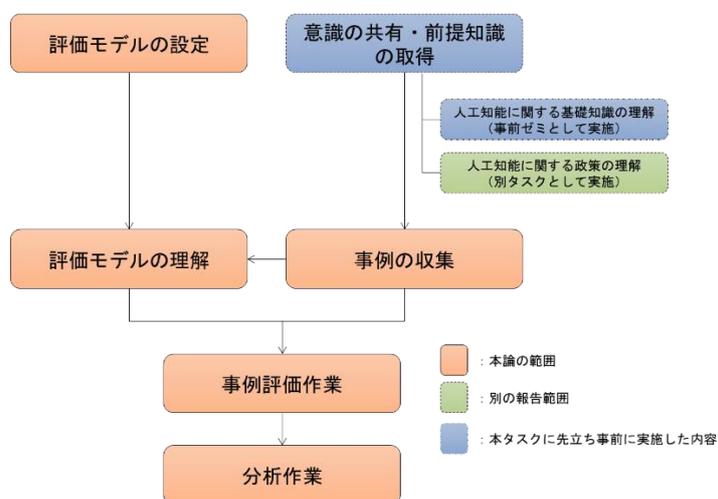


図 2 評価分析作業の流れ

実施した分析作業の全体をまとめると次表のようになる。多数の事例を対象にしたため、各事例の特性を評価する作業をメンバー全員で分担した。そしてその作業過程を共有し、判断の齟齬をきたさないように事前の意見交換と各評価結果のレビューを実施した。

表 4 本件評価分析作業の内容

作業番号	作業名	内容	実施時期
1000	意識の共有・前提知識の取得	本評価作業の実施に当たり必要となる第3世代の人工知能の概要や実際の取り組みについて、共通認識をはかる	本件着手前
1100	人工知能に関する基礎知識の理解 (事前ゼミとして実施)		本件着手前
1200	人工知能に関する政策の理解 (別タスクとして実施)		本件作業と 並行実施
2000	評価モデルの設定	本評価分析作業の目標を達成するために必要となる評価モデルを設定する	本件作業内
3000	評価モデルの理解	本件評価作業実施に当たり事例サンプルを用いた評価作業をとおして評価モデルの理解をはかる	本件作業内
4000	事例の収集	本評価作業のインプットとなる異なる3局面に関するデータの収集と、事前のデータクリーニング他分類作業の実施	本件作業内
5000	事例評価作業	本件作業メンバーによる前述3軸を用いた評価作業の実施	本件作業内
6000	分析作業	3軸による評価結果を用いた分析を行い、第三世代人工知能の限界を把握し、第3世代の人工知能の手法では処理できない領域を明確にするための検討のインプットを作成	本件作業内

2.1.4. 事例データの概要

(1) 日本経済新聞

最近の事例として、2017年中に日本経済新聞本誌に掲載された新聞記事の中で人工知能(AI)を見出しに含む記事を選定し、前述の分析モデルを用いた評価を実施した。

収集した新聞記事の概要は以下の通り。

表 5 新聞記事 事例の概要

新聞記事の対象期間	2017年1月1日～2017年12月31日
対象記事数	804件(確認できたもののみ)
対象条件	上記期間に含まれる記事のうち、見出しに人工知能(もしくはAI)を含むもの

これら収集した新聞記事をメンバー17名にて、分析モデルを用いた評価作業を実施しその内容を集計した。

(2) 人工知能学会誌

2017年に発行された、当該学会誌6号分について、所載論文と論説を中心に選定し、前述の分析モデルを用いた評価を実施した。

収集した学会誌に収録された論文の概要は以下の通り。

表 6 対象となる人工知能学会誌の概要

学会誌の対象期間	2017年に発行された6冊
対象となった論文・論説の件数	121件

これら収集した学会誌データをメンバーにより、分析モデルを用いた評価作業を実施しその内容を集計した。

(3) Preferred Networks 社⁴のホームページに開示された事例

Preferred Networks 社のホームページに開示された事例として、掲載当初(2014年9月28日)からの記事を選定した。これらの事例は深層学習フレームワーク Chainer を試用したもので、前述の分析モデルを用いて特性評価を実施した。

収集した公開事例記事の概要は以下の通り。

表 7 Preferred Networks 社事例の概要

事例記事の対象期間	2014年9月28日～2018年4月17日
対象記事数	71件
対象条件	当該企業のホームページ*に掲載されている全記事 * https://www.preferred-networks.jp/ja/news

これら収集したホームページ掲載のデータを所員メンバーにより、分析モデルを用いた評価作業を実施しその内容を集計した。

(4) 政府の取り組み事例 (NEDO)

NEDO「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」(調査研究)採択テーマに関し、公開された採択データ(平成27年度～平成29年度)からの事例を対象とし、前述の分析モデルを用いた評価を実施した。

⁴ Preferred Networks 社は2014年、3名の東京大学大学院生により設立されたスタートアップ企業で、深層学習フレームワーク Chainer を開発。世界454チームが参加した Google AI Open Images の物体検出コンペで準優勝した。Chainer はオープンソースとし、ディープラーニングを試行する多くのベンチャーや研究者に提供している。トヨタ、ファナック、パナソニック、中外製薬、東京エレクトロン等から資金調達を実現している。

収集した公開事例記事の概要は以下の通り。

表 8 NEDO 調査研究採択テーマにおける事例の概要

事例の対象期間	平成 27 年度～平成 29 年度
対象事例数	57 件（再委託、共同研究に細分化された単位）
対象条件	NEDO により採択結果情報*として公開された事例 * http://www.nedo.go.jp/content/100877631.pdf

これら収集したデータを所員メンバーにより、分析モデルを用いた特性評価作業を実施しその内容を集計した。

以下に、各データソースの具体的な内容について節を分けて説明する。

2.2. 日本経済新聞記事

日本経済新聞の対象記事については、本件分析作業に先立ち、記事内容から収集する情報の細分化を図るために、記事内容のクラスタリングを行い 22 の記事分類を付している。

なお、22 の記事分類の作成に当たっては、日本経済新聞社のホームページ(<https://www.nikkei.com/>)において、「人工知能 or AI or ディープラーニング or 深層学習」を検索キーワードに用いて検索を行った結果、2018 年 2 月 20 日時点で 11,109 件の記事がヒットし、その中でタイトル・サマリー部分がダウンロードできた 10,000 件から対象外となるものを除いた 8,242 件のデータを用いた。

初期段階で作成したクラスタについては、実際の作業を行う際の前段階で対象記事のクラスタへの振り分け状況を勘案し若干の見直しを行っている。

なお、作成したクラスタは以下の通りとなる。

表 9 日本経済新聞に掲載された人工知能関連記事の分類

No.	分類名	内容	記事数
1	AI スピーカー・会話・音声	AI スピーカーの提供や活用・関連技術、音声認識や AI と会話する技術関連	67
2	自動運転	自動運転に関する製品やサービスの提供・関連技術	32
3	囲碁・将棋	主にアルファ碁、DeepZenGO に関するもの、AI と人間との闘い	20
4	フィンテック（金融系含む）	金融サービスへの AI 活用や AI による変革、フィンテック	53
5	ロボット	ロボット（ロボ）に関する製品の提供・関連技術や AI による変革	31
6	働き方・業務効率化・業務支援	業務における人間の代替・省力化・効率化と AI、働き方改革と AI の関連	103
7	AI 人材	AI 関連技術者の不足、人材育成	27

No.	分類名	内容	記事数
8	ドローン	ドローンに関する製品の提供・関連技術	1
9	量子コンピュータ・スパコン	量子コンピュータ・AI 用スーパーコンピュータの研究・開発・関連技術	11
10	自動翻訳・通訳	自動翻訳・通訳に関する製品やサービスの提供・関連技術	7
11	コールセンター、自動応答、チャットボット	コールセンター、自動応答、チャットボットに関する製品やサービスの提供・関連技術や AI による変革	16
12	医療・創薬・ヘルスケア・介護	医療・創薬・ヘルスケア・介護に関する製品やサービスの提供・関連技術や AI による変革	48
13	物流・配送・配車	物流・配送・配車に関する製品やサービスの提供・関連技術や AI による変革	13
14	インフラ検査・点検	インフラ検査・点検に関する製品やサービスの提供・関連技術や AI による変革	37
15	AI 用半導体	AI 用半導体に関する動向、主要プレイヤーと半導体利用企業の関係性	9
16	スマート工場・自動工場	工場や生産現場分野に対する製品やサービスの提供・関連技術や AI による変革	14
17	農業・漁業	農業・漁業・畜産・農林水産分野に対する製品やサービスの提供・関連技術や AI による変革	9
18	防犯・防災	防犯や防災分野における AI の活用やその事例	8
19	生活	家電など生活分野において AI を含む製品やサービス関連	24
20	その他分野	上記に分類されない他の分野の製品やサービスといったもの	23
21	技術	AI に関する技術の解説やトレンドを取り扱ったもの（分野に特化していない）	53
22	トピック	論評やコラムなど AI 全般を含む記事や AI 関連企業の業績など	196

ちなみに、事前調査により収集した日本経済新聞社ホームページ検索によるクラスタ作成作業において抽出された記事タイトルとそのサマリに含まれる語句の分析結果は以下の通りとなっている。

なお、表 10 の事例数欄にあげた件数は、論文等の数ではなく、各論文から特徴評価により実際に抽出された事例数に相当する。

2.4. Preferred Networks 社の事例

当該事例データについては、事例記事内容から収集する情報の細分化を図るために、ホームページに掲載された事例記事のタイトルを元にクラスタリングを行い 6 つの記事群に分類した。

表 11 収集事例の分類

No.	分類名	内容	事例数
1	プラットフォーム（基盤技術）・スーパーコンピュータ・エッジコンピューティング	オープンソースの深層学習フレームワーク Chainer 関連製品やサービスの開発、Chainer を稼働させるインフラに関する事例	13
2	プロジェクト	他社との共同プロジェクトの事例	6
3	自動着色	自動着色製品の事例	2
4	インターン・事務所移転・その他	対象事例以外の記事（採用、インターン他）	8
5	協賛・出展・受賞	対象事例以外の記事（フェアや展示会への出展告知、社員の受賞・社外発表の告知）	25
6	企業間提携・資金調達・新会社設立	対象事例以外の記事（出資受け入れや提携に関する発表）	17

2.5. 政府の取り組み（NEDO 事例）

当該事例については、国主導の政策からの事例として、NEDO が実施している「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」に関する調査研究事業にてこれまで採択された案件を用いている。プロジェクトのねらいは、NEDO の「次世代人工知能・ロボット中核技術開発」基本計画によれば、「人工知能・ロボット関連技術の熟度に応じて、1) すでに技術的に確立し社会への普及促進が図られる段階、2) 技術的に概ね確立し実用化研究開発によりモデルを提示する段階、3) 人工知能・ロボットの利用分野を念頭におきつつ人間の能力を超えることを狙う、又は人間に匹敵する大きな汎用性、ロバスト性等を有する革新的な要素技術を研究開発する段階の三つの領域に整理する。本プロジェクトでは、単なる現在の人工知能・ロボット関連技術の延長上にとどまらない、人間の能力を超えることを狙う革新的な要素技術を研究開発する。」とされている。

また、次世代人工知能分野における研究開発にあたっては前述計画書によれば、4 つの研究開発項目に研究テーマを分類し、それぞれにおいて研究実施者を選定している。なお、当該事例については、平成 27 年度から平成 29 年度までの 3 カ年、公募にて採択されたものである。

本件事例評価作業においては、抽出した事例を NEDO が定めた 4 つの研究開発項目に用いた分類により整理している。

表 12 評価を実施した NEDO 採択案件の分類と抽出事例数

No.	研究開発項目	説明	採択年			事例数 合計
			2015	2016	2017	
1	①「大規模目的基礎研究・先端技術研究開発」	最新の計算論的神経科学の知見を取入れた脳型人工知能及びデータ駆動型的人工知能と知識駆動型的人工知能の融合を目指すデータ・知識融合型人工知能に関して、大規模なデータを用いた実世界の課題への適用とその結果の評価を前提とした目的基礎研究（大規模目的基礎研究）と、世界トップレベルの性能の達成を目指す先端技術の研究開発を実施する。	13	2	7	22
2	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」	広範な人工知能応用の研究開発や社会的実用化に資するため、研究開発項目①の成果である脳型人工知能技術、データ・知識融合型人工知能技術、その他大学や企業が保有する様々な人工知能技術をモジュール化し統合するための次世代人工知能フレームワークと、次世代人工知能技術を統合し、多様な応用に迅速につなげるための核となる先進中核モジュールの研究開発を実施する。	9	0	7	16
3	③「次世代人工知能共通基盤技術研究開発」	次世代人工知能の共通基盤技術として、人工知能技術の有効性や信頼性を定量的に評価し、性能を保証するための方法、そのために必要となる標準的問題設定や標準的ベンチマークデータセット等が満たすべき性質と構築の方法に関する研究開発を実施する。 また、それらを用いて、研究開発項目①、②の成果の評価を行う。	7	0	7	7
4	⑦「次世代人工知能技術の社会実装に関するグローバル研究開発」	次世代人工知能技術の社会実装が求められる領域として、「人工知能の研究開発目標と産業化のロードマップ」における当面の検討課題のうち、(1)生産性、(2)健康、医療・介護、(3)空間の移動の3領域において、関連する課題の解決に資するため、次世代人工知能技術の社会実装に関する研究開発を先導研究から実施する。 なお、人工知能技術とものづくり技術との融合等を国内外の叢智を結集して、グローバルに行うことを考慮する。	0	0	18	18
合計			29	2	39	70

※表中の説明は、「次世代人工知能・ロボット中核技術開発」基本計画（NEDO，2018年）より抜粋

※事例数は57件であるが、一つの事例で複数の領域を含むものがあるため70件となっている。

2.6. 政府の取り組み（全般）

政府全体のAI関連政策の取り組みについて、その全体像を本節にまとめる。

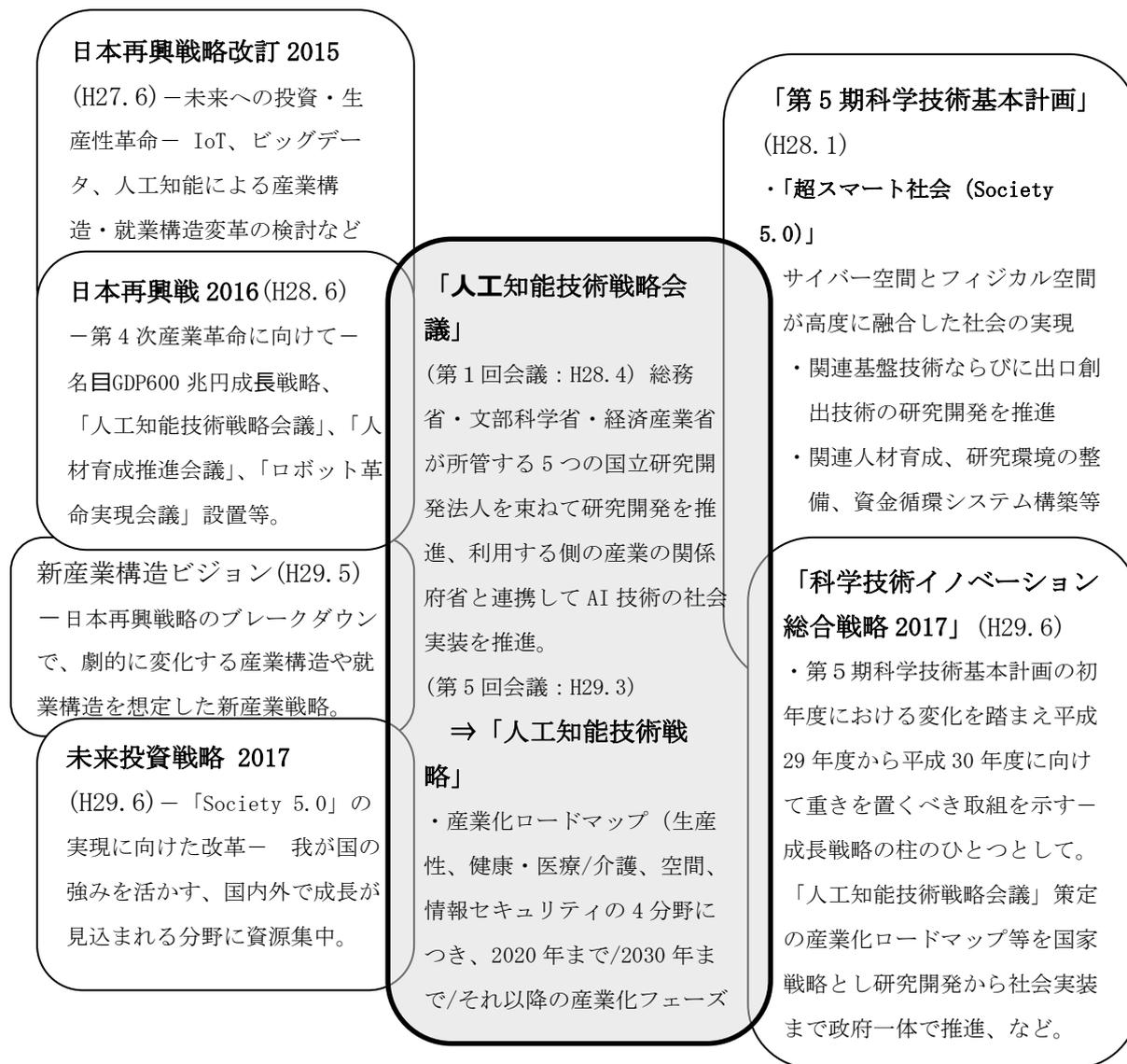
2.6.1. 調査分析の方法

ここ数年間の、わが国におけるAI関連政策動向については、NEDOの「AIポータルサイト」

(http://www.nedo.go.jp/activities/ZZJP2_100064.html) に主要な情報が集約されている。これによれば、「人工知能技術戦略会議」が関連の政策全体を束ねる機能を果たしていることが分かる。この会議体で約 1 年間の議論を経てとりまとめられた「人工知能技術戦略」、ならびに関連する上位政策を採りあげ、活動の実態や具体的内容を把握するとともに、次世代人工知能に関する取り組みの考え方を探った。

ここで取り上げた府省横断的な上位政策は以下に示す通りである。研究開発促進などの個別政策は、これら上位政策に記述された方針の範囲内で進められることから、ここでは採りあげていない。

＜調査の対象とした人工知能関連の上位政策群一時系列＞



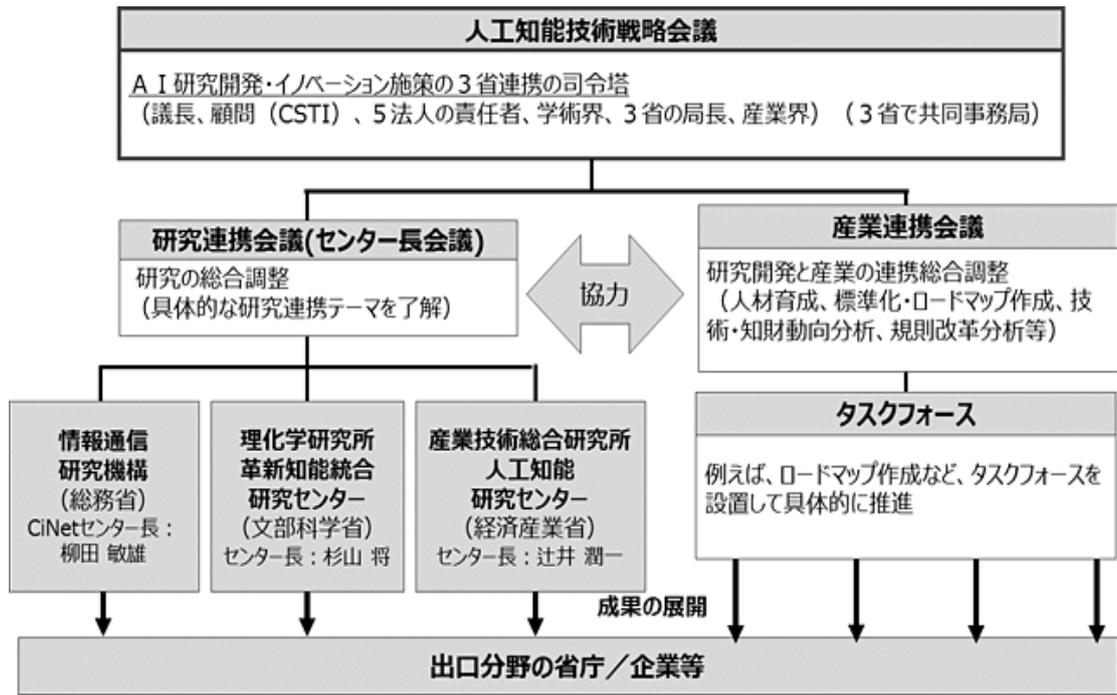
2.6.2. 「人工知能技術戦略会議」

(設立のいきさつ)

「日本再興戦略」改訂 2015 (平成 27 年 6 月 30 日閣議決定) では、IoT、ビッグデータ、人工知能による産業構造・就業構造変革の検討が主要施策の一つとして掲げられている。また、「第 5 期科学技術基本計画」(平成 28 年 1 月 22 日閣議決定)においても、「超スマート社会」の実現に向けた共通基盤技術や

人材の強化」として、AI等の重点的に取り組むべき技術課題等を、関係府省の連携の下で戦略的に研究開発を推進するとしている。

具体的な推進体制としては、平成28年4月12日に開催された第5回「未来投資に向けた官民対話」における安倍総理の指示を受け、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」が創設され、「日本再興戦略」2016（平成28年6月2日閣議決定）のなかに位置づけられている。同会議が司令塔となり、総務省・文部科学省・経済産業省が所管する「5つの国立研究開発法人」を束ね、人工知能（AI）技術の研究開発を進めるとともに、AIを利用する側の産業の関係府省と連携し、AI技術の社会実装を進めている。



・関係機関(上述の「5つの国立研究開発法人」):

- ①情報通信研究機構（NICT） 脳情報通信融合研究センター（CiNet） 知能科学融合研究開発推進センター
：自然言語処理、多言語音声翻訳や脳情報通信等の研究
- ②理化学研究所（RIKEN） 革新知能統合研究センター（AIP） 少ないデータからの高精度学習が可能となる新たなアルゴリズムなどの基礎研究・基盤技術の研究
- ③産業技術総合研究所（AIST） 人工知能研究センター（AIRC） ；それらの成果を活用し、ロボットの最適な動作を実現するなど産業分野への応用につなげる研究等
- ④科学技術振興機構（JST）
- ⑤NEDO

(会議体運営要領-20160418 制定、20160929 改)

総務省情報通信国際戦略局、文部科学省研究振興局、経済産業省産業技術環境局

- ・人工知能（AI）技術を核としたIoTの社会・ビジネスへの実装に向けた研究開発・実証につなげるため、総務省・文部科学省・経済産業省の3省が連携・合同して次世代のAI技術の研究開発と成果の社

会実装を加速するにあたり、AI 技術に関する技術戦略、3 省として実施すべき内容及びその役割分担、AI 技術に関する重要な事項について、関係機関の合議により処理することが適当な事務の方針等を調査・検討する会議。

- ・研究連携会議：研究連携会議は、次世代の AI 技術に関し、研究開発の総合調整を行うため、3 省が行う研究開発での連携の具体化等の調査・検討を行う会議。
- ・産業連携会議：産業連携会議は次世代の AI 技術に関し、研究開発と産業化の総合調整を行うため、3 省が行う研究開発その他の事業の社会実装の円滑化、加速化に係る施策等の調査・検討を行う会議。

- ・その他関連会議等

産業構造審議会 産業技術環境分科会（研究開発・イノベーション小委員会）

産業構造審議会 新産業構造部会

- ・3 省に加え、内閣府（戦略的イノベーション創造プログラム（S I P））、厚生労働省・国土交通省・農林水産省などビッグデータを有し出口産業を所管する府省でも AI 技術を活用したプロジェクトが企画されている。

- ・人工知能技術戦略会議議長及び構成員等

- ◎ 議長（運営要綱第 3 条第 1 項関係）

安西 祐一郎（独立行政法人日本学術振興会 理事長）

- 顧問（運営要綱第 4 条関係）

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（運営要綱第 3 条第 1 項関係）

内山田 竹志（日本経済団体連合会未来産業・技術委員会委員長）

小野寺 正（日本経済団体連合会未来産業・技術委員会委員長）

五神 真（国立大学法人東京大学総長）

坂内 正夫（国立研究開発法人情報通信研究機構理事長）

中鉢 良治（国立研究開発法人産業技術総合研究所理事長）

西尾 章治郎（国立大学法人大阪大学総長）

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）

松本 紘（国立研究開発法人理化学研究所理事長）

- 「人工知能技術戦略」（平成 29 年 3 月 31 日）の概要

<http://www.nedo.go.jp/content/100862413.pdf>)

(1)現状認識：現在進んでいる AI 技術は特定タスクを行う特化型の AI 技術であり、あくまで人間の能力を補完するもの。米国企業がプラットフォーム構築や社会実装面で主導、わが国のモノづくりへの応用や文化面のコンテンツ蓄積を今後どう活かせるかが課題。

(2)政策課題

①米中に劣後している研究開発投資促進

②データのデジタル化、個人情報保護や利用制約の仕組み、センサー等の情報入出力(伝送)環境整備など

社会実装の促進

③AI人材育成策（エンジニア・データサイエンティスト）

④大企業のベンチャーとの連携などオープンプラットフォームの形成とコスト低減の誘導

⑤処理速度を追求する学習、推論などのAIの用途に特化した半導体の開発利用、小型化・低消費電力化、脳型、量子などの全く新しいアーキテクチャの開発とデバイス・システム化、革新的ネットワーク（5G等）との組合せ、など。

(3)産業化ロードマップ

ー重点分野：社会課題として喫緊の解決の必要性、経済波及効果への貢献、AI技術による貢献の期待、の観点から下記4分野を特定。それぞれ、目指すべき社会像を設定したうえで、下記フェーズに沿った産業化のイメージを図示。（省略ーニーズ論的視点が中心であり、応用されるAI技術の具体的内容、水準を示す記述はない。）

ー産業化フェーズ：技術面から。社会実装課題解決にはさらなる時間が必要としている。

フェーズ1（各領域においてデータ駆動型のAI利活用が進む～2020年）

フェーズ2（個別の領域の枠を越えAI・データの一般利活用が進む～2025/2030年）

フェーズ3（各領域が複合的につながり合いエコシステムが構築される～2以降）

ー対象分野

①「生産性」

生産システム、物流を含むサービス産業の高度化・最適化により、ユーザー主導型のエコシステムが実現。創造力の増幅により、次々と新しいサービス・製品が生み出される社会が構築される。

②「健康、医療・介護」

医療・介護情報をビッグデータ化し、AIにより世界一の医療技術先進国、介護技術先進国、そして予防医療の高度化による健康長寿・高齢者就労社会、を構築する。

③「空間」

移動時間・空間における、作業、生活、娯楽を可能とする、全ての人に自由で安全な空間移動の確保する、移動手段の共有によるエコを実現する、バーチャル移動により新たな価値を創造する、など。

④「情報セキュリティ」

上記の分野を支える横断的分野と位置付け。信頼性、安定性に加えて、技術の秘匿性を重視すべきと認識。

(4)3センターを中心とした人工知能技術の研究開発及び社会実装に係る取組

ー研究開発

産業ロードマップに沿って重点化。基礎研究から社会実装まで一貫して取り組むべきもの、短期的な収益化が見込めず民間だけでは開発が進まないもの、国際標準化や共通基盤技術など協調領域であるものを採りあげる。3センターをハブとしつつ産学官連携によるオープンイノベーションの推進、出口産業を所管する関係府省との連携、民間投資の促進、などに留意する。

ー人材育成

フェーズ1ではトップレベルのAI人材（問題解決、具現化、実活用能力に優れる）の即戦力としての育成が急務、フェーズ2及び3ではAI技術が創造する価値を産業として普及させる人材育成が必要、処遇も含め研究環境を整備。内外のグローバル人材の活用、社会ニーズに沿った教育環境の整備に注力する。

ー産学官が有するデータ及びツール群の環境整備

3センターと関係府省が連携して、健康、医療、介護、交通、農林水産分野など社会ニーズにつながっているデータの活用、環境整備を進める。学習済みモデルの普及、産学官連携によるデータ整備・提供の強化、民間保有データの利活用促進などにも注力する。

ーベンチャー支援

既存の大企業とベンチャーの連携によるオープンイノベーションのプラットフォーム形成、目利き人材のネットワーク化、コーディネーター人材育成などを強化する。ベンチャー人材の育成、プレゼンズ段階での資金支援の強化も必要。

ーAI技術の開発に係る理解促進

AI技術の進化・普及に伴う社会への負の影響を克服し社会を豊かにすることについて、理解を促進する。AI技術の性能や安全性はアルゴリズムやデバイスだけでなく、使用するデータや環境に依存することから、製造者だけでなく、サービスの提供者、使用者にAI技術の理解の共有を進めることが必要。

(5)人工知能技術戦略のフォローアップ

「人工知能技術戦略会議」が定期的にフォローアップを行う。制度的な検討を要するものについては、「未来投資会議」などの検討の場に情報提供を行い、タイムリーな検討を促す。

(小括一「人工知能技術戦略」)

まず、人工知能技術の現状を、「特定タスクを行う特化型の人工知能技術であり、あくまで人間の能力を補完するもの」と認識している。

そうした人工知能関連の研究開発投資についてわが国は米国や中国に劣後していること、米国企業がプラットフォーム構築や社会実装面で主導していることなどを指摘。そうした現状を踏まえたわが国の課題は、人工知能のモノづくりへの応用や文化面のコンテンツ蓄積といった強みを活かすことであり、またエンジニアやデータサイエンティスト等の人材育成や、社会実装促進のための基盤整備を早急に進める必要があるとしている。

戦略内容を示す産業化ロードマップには、重点4分野ごとに目指すべき社会像を示したうえで、フェーズ（2020年まで、2025/2030年まで、それ以降）に沿った産業化イメージと取り組むべき課題を描いている。産学官の専門家の意見を集約したものとみられるが、ニーズ論的視点が中心であり、応用される人工知能技術の具体的な水準を示す記述はない。

全体として、応用すべき人工知能技術の発展段階的分析と今後の展開の想定、それへの戦略的取り組みの考え方などは明示されておらず、いわゆる第3世代人工知能を中心にその社会実装が世界的に進む中で、先ずはその流れに遅れないようにするための実践的戦略となっている。人工知能に関わる科学技術戦略の検討が急がれる。

2.6.3. 「日本再興戦略」の概要

以下に示すように、日本再興戦略 2016 で、名目 GDP600 兆円に向け「第4次産業革命の実現～IoT・ビッグデータ・AI・ロボット～【付加価値創出：30兆円(2020)】」が具体的に打ち出され、「人工知能技術戦略会議」、「第4次産業革命 人材育成推進会議」、「ロボット革命実現会議」などの設置を決定している。

それに続く、未来投資戦略 2017－「Society 5.0」の実現に向けた改革－（平成 29 年 6 月 9 日）は、日本再興戦略 2017 年版とでもいうべきものであるが、出口を目指した実装戦略となっている。その重点推進 5 分野ならびに共通基盤強化方針は、「人工知能戦略」と整合しているが、逆に、新しい構想や戦略が示されているわけではない。

<https://www.kantei.go.jp/jp/singi/keizaisaisei/kettei.html>

（日本再興戦略の流れ）

○未来投資戦略 2017－「Society 5.0」の実現に向けた改革－ 平成 29 年 6 月 9 日

- ・実装戦略－出口論：我が国の強み（モノづくりの強さ、社会課題の先進性・大きさ、「リアルデータ」の取得・活用の可能性）を活かす、国内外で成長が見込まれる、などの視点から、①健康寿命の延伸、②移動革命の実現、③サプライチェーンの次世代化、④快適なインフラ・まちづくり、⑤ FinTech、の 5 分野を中心に政策資源を集中投入し、未来投資を促進するとともに、オープンな「データ基盤（リアルデータプラットフォーム）」の構築や IT 人材育成など共通基盤の強化を図る、などとしている。

* 「リアルデータ」については、「新産業構造ビジョン」の項を参照。

○日本再興戦略 2016－第4次産業革命に向けて－ 平成 28 年 6 月 2 日

- ・名目GDP600兆円に向けた成長戦略
- ・第4次産業革命の実現～IoT・ビッグデータ・AI・ロボット～
【付加価値創出：30兆円(2020)】
- ・政府全体の司令塔として「第4次産業革命官民会議」を設置。
（「未来投資に向けた官民対話」の機能を事実上置き換え）
- ・同会議の下に、「人工知能技術戦略会議」
「第4次産業革命 人材育成推進会議」
「ロボット革命実現会議」
を位置づけ、日本として「取るべき」重点分野の特定、「重点分野別戦略」の策定、横断的施策（規制改革、研究開発、資金供給、人材育成等）の加速化を図る。

○日本再興戦略改訂 2015－未来への投資・生産性革命－ 平成 27 年 6 月 3 0 日

- ・アベノミクス第二ステージ成長戦略： 「未来投資による生産性革命の実現」と「ローカル・アベノミクスの推進」を車の両輪とするもの。前者について、IoT・ビッグデータ・AI・ロボット等を活用している。

2.6.4. 第5期科学技術基本計画（平成28年1月22日閣議決定）

「日本再興戦略2016」に先行して、「超スマート社会（Society 5.0）」の実現を目指すことが打ち出されている。「Society 5.0」は、狩猟、農耕、工業、情報に次ぐ社会と位置付け、第4次産業革命におけるIoT、ビッグデータ、AI、ロボット等による技術革新により、「必要なもの・サービスを、必要な人に、必要な時に、必要なだけ提供し、社会の様々なニーズにきめ細かに対応でき、あらゆる人が質の高いサービスを受けられ、年齢、性別、地域、言語といった様々な違いを乗り越え、生き活きと快適に暮らすことのできる社会」の実現が可能とし、それを「超スマート社会」と表現。

時代の潮流として、「価値観の変化」、「知識・価値の創造プロセスの変化－オープンイノベーションの重視、オープンサイエンスの潮流」を捉え、「非連続なイノベーションを生み出す研究開発と、新しい価値やサービスが次々と創出される超スマート社会」を世界に先駆けて実現するための仕組みづくりを強化する、としている。

その「超スマート社会」を、「サイバー空間とフィジカル空間（現実社会）が高度に融合した社会」と定義し、それを支える基盤技術として、サイバーセキュリティ、IoTシステム構築、ビッグデータ解析、AI、デバイス等を、出口を創出するロボット、センサー、バイオ、素材・ナノ、光・量子などに関わる技術を挙げている。

目指す社会について、狩猟、農耕、工業、情報に次ぐ社会としての、期を画する中核的な概念ならびにそれへの変革動因は示されておらず、その実現に必要な「外部装置」としての科学技術ならびにそれを付加価値に転化する社会的な仕組みと多様な「補助装置」とそれへの科学技術の寄与など、科学技術政策の起点が不明である（「外部装置」や「補助装置」の意味については、「科学技術を契機とする我が国未来社会形成のための政策的対応に関する調査研究－社会的基盤形成と社会的受容に係る事例分析を手掛かりにして－」平成28年12月公益財団法人未来工学研究所を参照）。

開発対象あるいは実現を目指す人工知能のレベルは明確ではないが、上記の定義からは、自然システムや人工的物理システムならびに人工的抽象システムを対象とした「第3世代型人工知能」のレベルであり、「人間活動システム」を対象としているものではないと考えられる。並行して進められる人工知能の自主研究のなかには、「人間活動システム」を対象とするものが含まれる可能性はある。

（概要） <http://www8.cao.go.jp/cstp/kihonkeikaku/5honbun.pdf>

・ICTの進化等により、社会・経済の構造が日々大きく変化する「大変革時代」が到来し、国内外の課題が増大、複雑化する中で科学技術イノベーション推進の必要性が増していると認識。

・目指すべき国の姿として以下を掲げ、

- ①持続的な成長と地域社会の自律的発展
- ②国及び国民の安全・安心の確保と豊かで質の高い生活の実現
- ③地球規模課題への対応と世界の発展への貢献
- ④知の資産の持続的創出

・その実現のために4つの柱を推進し、「世界で最もイノベーションに適した国」へと導いていくこととしている。

- i) 未来の産業創造と社会変革

自ら大きな変化を起こし、大変革時代を先導していくため、非連続なイノベーションを生み出す研究開発を強化し、新しい価値やサービスが次々と創出される「超スマート社会」を世界に先駆けて実現するための一連の取組を更に深化させつつ「Society 5.0」として強力に推進する。

ii) 経済・社会的な課題への対応

国内又は地球規模で顕在化している課題に先手を打って対応するため、国が重要な政策課題を設定し、課題解決に向けた科学技術イノベーションの取組を進める。

iii) 基盤的な力の強化

今後起こり得る様々な変化に対して柔軟かつ的確に対応するため、若手人材の育成・活躍促進と大学の改革・機能強化を中心に、基盤的な力の抜本的強化に向けた取組を進める。

iv) 人材、知、資金の好循環システムの構築

国内外の人材、知、資金を活用し、新しい価値の創出とその社会実装を迅速に進めるため、企業、大学、公的研究機関の本格的連携とベンチャー企業の創出強化等を通じて、人材、知、資金があらゆる壁を乗り越え循環し、イノベーションが生み出されるシステム構築を進める。

2.6.5. 科学技術イノベーション総合戦略 2017（平成 29 年 6 月 2 日 閣議決定）

科学技術イノベーションは成長戦略の重要な柱と位置付けられており、総合科学技術・イノベーション会議が司令塔となり、「科学技術イノベーション総合戦略」を毎年策定し、施策の重点化等を実行することとなっている。科学技術イノベーション総合戦略 2017 は、第 5 期科学技術基本計画の初年度における変化を踏まえ、同計画を基に、平成 29 年度から平成 30 年度に向けて重きを置くべき取組を示している。

以下に項目的概要を示すが、この第 2 章に、「プラットフォームの構築」として、新たな価値やサービスの創出の基となるデータベースの構築と利活用、基盤技術としてのサイバー空間関連基盤技術の強化（AI 技術、ネットワーク技術、ビッグデータ解析技術等）、人工知能技術戦略会議策定の産業化ロードマップ等を国家戦略とし研究開発から社会実装まで政府一体で推進、フィジカル空間関連基盤技術の強化（ロボット技術、デバイス技術、素材・ナノテクノロジー、光・量子技術等）などが取り上げられているが、既に上述の上位政策で示されている内容に準じた項目の提示であり、目標とする技術レベルの表現はない。

（科学技術イノベーション総合戦略 2017 のポイント）

<http://www8.cao.go.jp/cstp/sogosenryaku/2017.html>

第 1 章 重点事項

世界に先駆けて Society 5.0 を実現するための、「科学技術イノベーション官民投資拡大イニシアティブ」の着実な実行、「Society 5.0 の推進と政府研究開発投資目標の達成に向けて」の着実な実行を掲げるとともに、卓越した研究拠点や多様な学術研究を生み出す研究環境を確保することが重要としている。

第 2 章 未来の産業創造と社会変革に向けた新たな価値創出の取組

未来に果敢に挑戦する研究開発と人材の強化、新たな社会としての「Society 5.0」を実現するプラットフォームの構築など

第3章 経済・社会的課題への対応

持続的な成長と地域社会の自律的な発展、国及び国民の安全・安心の確保と豊かで質の高い生活の実現、地球規模課題への対応と世界の発展への貢献、国家戦略上重要なフロンティアの開拓など

第4章 科学技術イノベーションの基盤的な力の強化

人材力の強化、知の基盤の強化、資金改革の強化など

第5章 イノベーション創出に向けた人材、知、資金の好循環システムの構築

オープンイノベーションを推進する仕組みの強化、新規事業に挑戦する中小・ベンチャー企業の創出強化、イノベーション創出に向けた知的財産・標準化戦略及び制度の見直しの整備、「地方創生」に資するイノベーションシステムの構築、グローバルなニーズを先取りしたイノベーション創出機会の開拓など

第6章 科学技術イノベーションの推進機能の強化

大学や国立研究開発法人の聖域なき改革の断行、研究開発の特性を踏まえた政府調達手続の迅速化、エビデンスに基づく実効性ある科学技術イノベーション政策の推進と総合科学技術・イノベーション会議の司令塔強化

2.6.6. 新産業構造ビジョン（平成29年5月30日 産業構造審議会新産業部会事務局）

「日本再興戦略」の産業政策としてのブレークダウンであり、関連する上位政策を踏まえ、2020年から2030年にかけて概ね見通せる現実的政策論である。すなわち、成長戦略第2ステージの課題のなかで最大の鍵は第4次産業革命技術の社会実装であり、IoT、ビッグデータ、AI、ロボットなどにより、これまで実現不可能と思われていた社会「Society 5.0」の実現が可能になるとする。これに伴い、産業構造や就業構造が劇的に変わる可能性があり、AIの非連続的進化等を踏まえた新産業戦略が必要としている。

戦略のポイントのひとつが「リアルデータのプラットフォームの創出」。圧倒的に出遅れている「検索エンジンのデータ競争」は困難であるが、健康・医療・介護、製造現場、自動走行等、日本にはモノづくりの延長にある、或いは地域独自の「リアルデータ」の蓄積があり、産業の枠を超え、企業横断的なプラットフォームを構築できれば、強みを発揮できるはず、としている。GAF(A Google, Apple, Facebook, Amazon)などが吸い上げてきた膨大な顧客等情報とは異なるビッグデータの活用である。

ここでも、適用すべき人工知能の技術レベルについては触れていない。

www.meti.go.jp/committee/sankoushin/shin_sangyoukouzou/pdf/017_05_00.pdf

(概要) ～副題：「1人ひとりの、世界の課題を解決する日本の未来」

・目指すべき将来像を「第5期科学技術基本計画」に示された「Society5.0」とし、それに対応する産業を「Connected Industries」と表現、「多様な人、組織、機械、技術、国家がつながり、新たな付加価値を創出し、社会課題を解決していくもの」と位置付けている。

・解決すべき課題、ニーズに対応する戦略4分野を掲げている。

①「移動する」(人の移動、モノの移動)

②「生み出す・手に入れる」(スマートサプライチェーン、製造・生産現場における高度化・効率化)

- ③「健康を維持する・生涯活躍する」(健康・医療・介護)
- ④「暮らす」(「新たな街」づくり、シェアリング、Fintech)

・対応する産業構造・就業構造の変革に向けて、新たな経済社会システム構築に向けた横断的施策を提示している。

- ①ルール的高度化(データ、知財・標準、規制制度改革)
- ②人材育成・活用システム(人材、教育、雇用)
- ③イノベーションエコシステム(CoE、大学、ベンチャー)
- ④経済の新陳代謝システム(ガバナンス、資金、産業再編)
- ⑤地域・中小企業システム(地域未来投資促進法、IT導入支援)
- ⑥社会保障システム(マイナンバー活用による個別化等)

2.6.7. 「人工知能と人間社会に関する懇談会」報告書(170324 内閣府←160530)

第5期科学技術基本計画で掲げた Society 5.0 の実現の鍵である人工知能の研究開発及び利活用を健全に進展させるべく、人工知能と人間社会の関わりについて検討を行うため、内閣府特命担当大臣(科学技術政策)の下に開催された「人工知能と人間社会に関する懇談会」(第1回平成28年5月30日開催(160530))において報告書がとりまとめられた。

<http://www8.cao.go.jp/cstp/tyousakai/ai/summary/index.html>

(現状認識)

人間の知的活動(認知、思考、推論、それらに基づく行為等)を部分的に代替し得る人工知能技術(以下、AIと略)は、ビッグデータと機械学習により急速に発展、自動運転や医療診断支援、対話エージェントなどへの実装も進んでいる。日本政府は、Society5.0の重要な基盤技術と位置付けている。

(報告書趣旨)

見えにくい技術であり、知らぬ間に社会のあり方に根本的影響を与える可能性があり、健全な利用に向けて影響の検討が必要。国内外でAIが社会に及ぼす影響(ELSI)の検討が進んでいる。

現存する、あるいは近い将来実現する可能性が高いAIが普及した社会に焦点を当て、期待される便益と考慮すべき点、今後の課題や方向性を明らかにすることを目的としている。移動、製造、個人向けサービス、対話・交流という4つの代表的分野について事例検討。国際的な協調により、国際的な枠組みの検討や共有できる価値観の追求が必要。

*倫理的論点: AIに基づく判断と人の判断とのバランスを取ることが重要。利用者が知らぬ間に感情や心情、行動が操作されたり、順位付け・選別されたりすることが生じ得る場合には倫理的検討が必要。人間観、倫理観、価値観の変化にも留意が必要。

*法的論点: AI活用時とそれを回避した場合の両方のリスクを考慮し、AIの利活用に伴う法的課題を抽出、たとえば事故時の責任分配や保険の整備、プライバシー保護や創作物権利保護、などを検討。社会の枠組み変化に対する基礎研究が必要。

*経済的論点: 仕事の質が変化することへの個人ならびに企業の対応が必要であり、政府は教育や環境の整備などが必要。

*社会的論点：社会コストの不均衡や差別の増加などについて、多様性や共通性の観点から自由に対話し検討することが必要。AIへの依存や化新、過剰な拒絶などの社会的病理が生じる可能性も検討すべき。

*研究開発的論点：倫理規定やガイドラインに沿った行動が求められる。研究内容的については、セキュリティ確保や制御不能回避、計算過程や論理の透明性確保などが重要。また、社会受容性を重視し、オープンサイエンスの促進、人文社会科学研究や融合研究の推進に留意する必要がある。

2.6.8. まとめ

ここ数年間の、わが国における人工知能関連の上位政策群の概要ならびに相互関係などの実態を把握するとともに、次世代人工知能に関する取り組みの考え方を探った。結果として以下のことが云える。

①人工知能関連政策については、ここ数年で主要な上位政策群が出揃った。「日本再興戦略 2016」に位置づけられる「人工知能戦略会議」の設置（2016年）や再興戦略の産剛政策としての「新産業構造ビジョン」（2017年）、これらに先立つ「第5期科学技術基本計画」（2016年）およびその実行戦略である「科学技術イノベーション総合戦略 2017」など。

②これら主要政策を踏まえた公的研究の推進は、「人工知能戦略会議」に参画する「5つの国立研究開発法人」を中心に進められることとなり、個別に進められていた研究全体がより戦略的に再編・再構築された。

③人工知能の研究開発や社会実装に関わるこれら上位政策に共通する内容は、やや出遅れた感のある人工知能実装に関わる内容が中心となっている。従って、ニーズ論的視点が中心であり、応用される人工知能技術の具体的な水準を示す記述はない。すなわち、応用すべき人工知能技術の現状に対する発展段階的分析と今後の展開の可能性、それへの戦略的取り組みの考え方などは明示されていない。いわゆる第3世代人工知能を中心にその社会実装が世界的に進む中で、先ずはその流れに遅れないようにするための実践的戦略となっている。

④人工知能に関わる科学技術戦略の検討が急がれる。上述の政策群には、大学との連携がうたわれているものの、次世代人工知能に関わる研究戦略構築には至っていない。関連のシンポジウムを垣間見ると、研究専門家の人工知能に関する現状の認識と将来の進むべき方向についての様々な意見が交わされている。公的研究機関と主要大学だけでなく、内外の各大学・研究機関の独自の研究動向にも注目する必要がある。こうした広範囲な研究内容に関わる動向の調査・分析には、別途取り組みが必要となる。

⑤内閣府所管の「人工知能と人間社会に関する懇談会」では、人工知能の研究開発及び利活用を健全に進展させるべく、人工知能が社会に及ぼす影響（ELSI）の検討を行っている（2017年3月に報告書）。時宜を得た研究であり、今後とも世界と連携しつつ政策的に研究を継続することが必要と考えられる。

3. 得られた知見のまとめ

3.1. 評価分析結果のサマリについて

先にあげた収集事例について、前述の3軸による評価を行い、評価結果を用いて以下の項目により分析を行った。

- 1) 全体サマリ
- 2) 事例パターン
- 3) 各軸間のクロス集計
- 4) その他

3.1.1. 全体サマリ

全体サマリは、4つの事例ソースにおいて、それぞれ3つの評価軸単位にどの程度の事例数が登場するかを俯瞰したものである。なお、記事・論文に含まれる事例については、一記事一事例とはならず、一記事に複数事例を含むものがあるため、それらを個別の事例数としてカウントしている。

なお、「人工知能を作動させている対象の知識カテゴリ」や「人工知能手法の発展段階レベル」が特定できないもの（言及されていないなど）も存在する。

次にあげる表中の各調査対象については、色の濃い部分がそれぞれの対象のカテゴリ中に多く登場する事例を含む内容となっている。

各データソース別に全体像を以下に示す。

① 日本経済新聞

No.	記事カテゴリ名	記事数	事例件数		人工知能のレベル					対象の知識カテゴリ						ステージ						
			対象	対象外	レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	1	2	3	4	5	6	不明
1	AIスピーカー・会話・音声	67	90	26	0	3	6	35	0	0	1	1	43	0	0	0	0	6	46	1	1	62
2	自動運転	32	32	0	0	0	0	7	3	0	0	0	28	0	4	1	2	23	2	2	2	5
3	囲碁・将棋	20	18	2	0	0	6	17	0	0	3	4	10	0	0	0	1	9	1	6	0	3
4	フィンテック(金融系含む)	53	53	0	5	17	37	28	0	0	3	0	30	35	0	3	2	1	12	32	1	2
5	ロボット	31	13	18	0	2	9	3	0	5	6	2	11	0	0	0	3	7	3	0	0	18
6	働き方・業務効率化・業務支援	103	107	5	2	18	53	24	8	15	10	16	54	9	8	2	5	24	43	25	1	14
7	AI人材	27	22	5	0	0	2	3	9	0	1	0	5	1	9	0	0	0	1	4	7	15
8	ドローン	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9	量子コンピュータ・スパコン	11	10	1	0	1	2	6	0	10	0	0	0	0	0	2	3	3	2	2	2	2
10	自動翻訳・通訳	7	9	0	0	1	5	4	0	0	0	0	9	0	0	0	0	0	2	5	2	0
11	コールセンター、自動応答、チャットボット	16	16	0	0	3	10	3	3	0	0	0	16	0	0	0	0	4	5	8	0	1
12	医療・創薬・ヘルスケア・介護	48	44	4	0	0	23	9	0	0	0	0	36	0	0	0	14	4	1	10	12	8
13	物流・配送・配車	13	14	0	2	1	11	5	1	0	0	8	9	3	0	0	2	7	8	1	0	0
14	インフラ検査・点検	37	36	1	0	10	22	4	0	34	5	0	0	0	0	2	15	15	10	6	5	1
15	AI用半導体	9	5	4	0	0	2	3	0	1	0	2	2	2	0	1	4	1	1	0	0	0
16	スマート工場・自動工場	14	15	0	0	1	5	9	1	6	3	1	6	0	1	0	0	4	8	3	0	0
17	農業・漁業	9	15	0	6	2	3	0	2	7	3	0	0	0	2	0	0	3	7	0	0	5
18	防犯・防災	8	8	0	1	0	6	4	1	1	1	2	5	0	1	0	0	2	5	0	0	1
19	生活	24	20	5	0	5	11	4	4	0	1	7	9	3	1	0	2	3	9	1	2	8
20	その他分野	23	24	0	0	0	4	2	0	2	0	0	16	0	1	2	2	6	7	0	7	0
21	技術	53	39	14	0	8	18	16	14	3	4	11	20	4	11	0	4	9	11	3	4	26
22	トピック	196	186	23	13	20	45	50	68	10	7	9	68	31	48	7	19	15	41	32	39	73
	合計	802	776	109	29	92	280	236	114	94	48	63	375	88	86	20	78	146	225	141	85	245

② 人工知能学会誌

号数	特集・テーマ	抽出事例数		人工知能のレベル					対象の知識カテゴリー					ステージ							
		対象	対象外	レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	1	2	3	4	5	6	不明
1	1. 編集委員今年の抱負	36	16	0	2	7	16	0	3	0	1	9	10	1	13	3	4	2	2	1	17
	2. Well-being Computing	7	0	0	0	2	0	0	0	0	0	5	1	0	0	0	0	3	0	2	1
2	1. ゲーム産業における人工知能	5	0	0	2	3	2	1	0	0	4	1	1	0	0	0	0	4	0	1	0
	2. 身体知の発展	8	0	0	1	2	1	1	0	0	3	1	4	0	1	6	0	0	0	1	0
	3. 日本のAI: 元気な若手の動き	7	1	0	0	2	4	2	1	0	2	0	2	0	5	5	1	1	0	0	0
3	0. 巻頭言	3	0	0	0	3	0	0	1	0	0	0	1	0	0	3	3	3	0	0	0
	1. 顔文字の科学	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2. 人文科学とAI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1. 広告とAI	8	0	0	0	8	0	0	2	0	4	2	0	0	1	3	0	3	1	0	0
5	2. 不動産とAI	12	3	0	0	6	4	0	0	2	7	0	0	5	5	2	0	2	0	0	0
	1. AI社会論	9	0	0	0	1	7	6	2	5	2	4	6	0	3	3	1	0	0	0	2
6	2. 人と調和して協働する知的情報処理	7	2	0	0	0	4	1	0	3	0	4	0	0	0	0	5	0	0	0	2
	3. 人間を超えるコンピュータ囲碁	7	1	0	0	6	6	0	0	0	6	1	0	0	0	0	0	6	0	1	0
	1. 脳科学とAIのフロンティア	7	1	0	0	0	6	0	0	6	0	0	0	3	7	0	0	0	0	0	0
7	2. マスメディアから見た人工知能	5	1	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	1	1	0
	3. 全国大会(企画+OS)	27	7	0	4	17	13	0	9	0	10	6	2	14	8	8	1	1	1	8	0
合計		148	32	0	9	57	63	11	16	10	20	51	34	3	45	43	27	14	14	8	32

③ Preferred Networks 社のホームページに開示された事例

No.	記事分類	記事数	抽出事例数		人工知能のレベル					対象の知識カテゴリー					ステージ							
			対象	対象外	レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	1	2	3	4	5	6	不明
1	プラットフォーム・スーパーコンピュータ・エッジコンピューティング	13	12	1	0	0	0	9	3	1	0	7	1	0	3	0	0	0	3	7	9	0
2	プロジェクト	6	7	1	0	0	3	3	1	5	0	5	1	0	1	0	0	1	6	1	0	0
3	自動着色	2	2	0	0	0	0	2	0	0	0	2	0	0	0	0	0	0	2	0	0	0
4	インタビュー・事務所移転・その他	8	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	協賛・出展・受賞	25	0	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	企業間提携・資金調達・新会社設立	17	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
合計		71	21	52	0	0	3	14	4	6	0	12	4	0	4	1	2	4	15	13	15	0

④ 政府の取り組み事例 (NEDO)

No.	研究開発項目	事例数	人工知能のレベル					対象の知識カテゴリー					ステージ								
			レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	1	2	3	4	5	6	不明	
1	①「大規模目的基礎研究・先端技術研究開発」	22	0	0	0	0	22	0	14	5	0	13	2	0	3	8	6	5	0	0	0
2	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」	16	0	0	1	16	0	4	9	4	6	1	0	1	4	6	5	0	0	0	0
3	③「次世代人工知能共通基盤技術研究開発」	14	0	0	0	14	0	3	7	1	3	1	0	0	7	2	5	0	0	0	0
4	④「次世代人工知能技術の社会実装に関するグローバル研究開発」	18	0	0	0	18	0	8	10	0	0	0	0	0	7	10	1	0	0	0	0
合計		70	0	0	1	70	0	29	31	5	22	4	0	4	26	24	16	0	0	0	0

3.1.2. 事例パターン

事例パターンとは抽出された各事例について、たとえば「人工知能手法の発展段階レベル」と「人工知能を作動させている対象の知識カテゴリー」を構成する各項目の組み合わせを、同じ組み合わせとなるもの同士（パターンと呼ぶ）を集計し一覧に示したものである。

この結果から、たとえば新聞記事で多く登場する「人工知能手法の発展段階レベル」と「対象となる知識のカテゴリー」の組み合わせについては、「レベル4 & ソフト2」、もしくは「レベル3 & ソフト2」であることを見て取ることができる。

下記に挙げたパターンのうち、「発展段階レベル」、「知識カテゴリー」双方が一意にきまるケースを見ると、大半はレベル3・ソフト2、レベル4・ソフト2の事例パターンが多く登場する(Preferred Networks社の事例を除く)。

ソフト2は、「境界領域もしくは人間・社会システム」に対する「人工的抽象システム」である。ピンクの網掛けで指定するソフト3は、「境界領域もしくは人間・社会システム」を対象とするが、「人間活動システム」であり、「発展段階レベル」が3、4であったとしても、「人間活動システム」を実現する人工

知能の事例は多くないことがわかる。

① 日本経済新聞（数の多いもののみ）

分類No.	事例数	人工知能のレベル					対象の知識カテゴリー						
		レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	
1	91			○						○			
2	84				○					○			
3	70					○						○	
4	64	人工知能のレベルが特定されないもの								○			
5	24			○	○					○			
6	22					○				○			
7	20			○			○						
8	15				○		○						
9	14		○							○			
10	13			○				○					
11	13	人工知能のレベルが特定されないもの					○						
12	11		○	○						○			
13	10				○						○		
14	10				○		対象の知識カテゴリーが特定されないもの						
15	10			○					○				
16	9					○					○		
17	8	○	○	○	○						○		
18	8			○	○					○	○		
19	7			○	○		○						
20	7				○					○	○		
21	7				○				○				
22	7		○	○			t			○	○		
23	7	○					○						
24	6			○	○				○				
25	6				○			○					
26	6		○	○			○						
27	6			○					○	○			

② 人工知能学会誌

分類No.	事例数	人工知能のレベル					対象の知識カテゴリー						
		レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	
1	11			○						○			
2	9			○	○					○			
3	8				○					○			
4	7				○						○		
5	7				○				○				
6	4			○	○					○	○		
7	4			○	○		○						
8	4			○					○				
9	4										○		
10	4	人工知能のレベルが特定されないもの										○	
11	3			○	○		○			○			
12	3			○			○						
13	3			○							○		
14	3	人工知能のレベルが特定されないもの							○				
15	2			○	○							○	
16	2			○	○	○					○		
17	2				○					○	○		
18	2				○			○					
19	2		○	○			○						
20	2			○				○					
21	2					○					○		
22	1		○	○	○				○				
23	1		○	○	○						○		
24	1			○	○	○				○			
25	1			○	○	○	対象の知識カテゴリーが特定されないもの						
26	1			○	○				○				
27	1			○	○						○		
28	1			○	○	○	○	○	○	○	○		
29	1			○	○	○			○	○	○		
30	1			○	○			○		○	○		
31	1			○	○			○		○	○		
32	1				○			○			○		
33	1				○		○						
34	1				○							○	
35	1		○	○							○		
36	1		○	○							○		
37	1			○			対象の知識カテゴリーが特定されないもの						
38	1			○					○	○			
39	1		○				○						
40	1		○						○				
41	1		○							○			
42	1					○	○	○	○	○	○		
43	1						○		○				
44	38	対象外、もしくは不明などの理由で分類が付されていないもの											

なお、ソフト3を含むケースの多くは、「人工知能のレベルが特定できない」・「その他」に該当する事例が多くなっている。

また、対象の知識カテゴリの「自然・工学システム」に該当するハード1やソフト1を含むパターンにおいては、「発展段階レベル」1・2だけでなくレベル3・4（特に、4のほうが多い）も数多く登場する。

③ Preferred Networks 社のホームページに開示された事例

分類No.	事例数	人工知能のレベル					対象の知識カテゴリ					
		レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他
1	6				○				○			
2	4					○						○
3	4				○					○		
4	3				○		○		○			
5	3			○			○		○			
6	1				○							
7	52	いずれにも該当しない事例(対象外)										

④ 政府の取り組み事例（NEDO）

分類No.	事例数	人工知能のレベル					対象の知識カテゴリ					
		レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他
1	21				○			○				
2	14				○		○					
3	11				○		○			○		
4	3				○		○		○	○		
5	2				○				○	○		
6	2				○					○		
7	1				○		○				○	
8	1				○			○	○	○		
9	1			○	○			○				
10	1				○						○	

3.1.3. 各軸間のクロス集計

3つの軸について、以下のクロス集計を行った。

- (1) 「発展段階レベル」と「知識カテゴリ」
- (2) 「発展段階レベル」と「研究開発ステージ」
- (3) 「知識カテゴリ」と「研究開発ステージ」

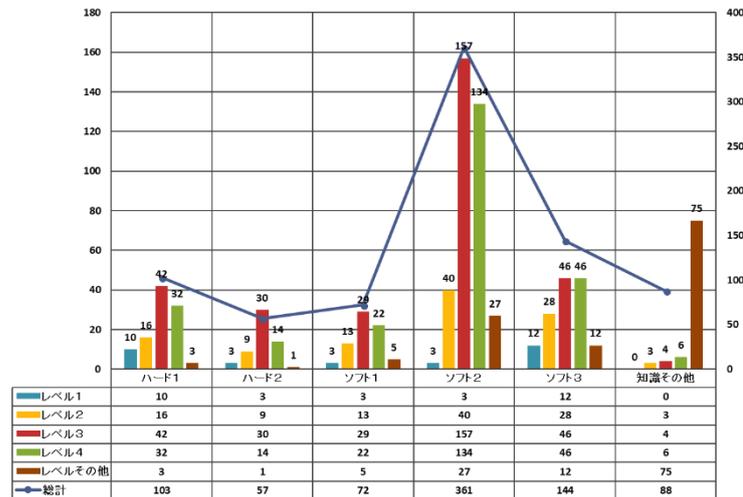
(1) 「発展段階レベル」と「知識カテゴリ」

異なる4データソースの事例において、1:「① 日本経済新聞と② 人工知能学会誌」、2:「③ Preferred Networks 社と④ 政府の取り組み事例 (NEDO)」の2つの傾向があることが見て取れる。2:は、先端の人工知能関連ベンチャーと NEDO で取り上げられるテーマが似通っていると考えることができる。両社とも、人工知能のレベルは4であるが、自然工学システムを対象とした技術開発に向かっているとみることができる。

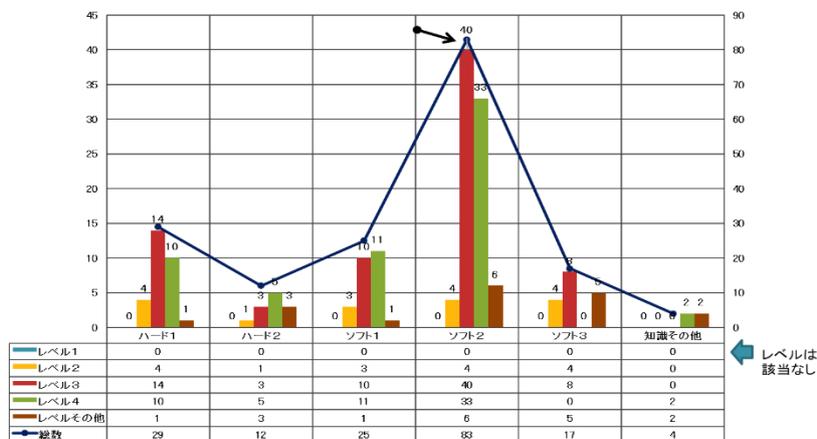
また、1:では、ソフト2を中心に「発展段階レベル」が3ないし4の事例が多く、「人間・社会システム」を対象としたものがより多く取り上げられる傾向にあると考えられる。特に、② 人工知能学会誌では、ソフト2の事例数が多いことから、より人間・社会システムを対象とした事例が多くなっている。

人間活動システムを対象とした事例や、研究の方向性を見るにあたっては、1:のようなデータソースの事例をさらに深堀する必要があると考えられる。②人工知能学会誌では、レベル4をソフト3に適用した例はない。

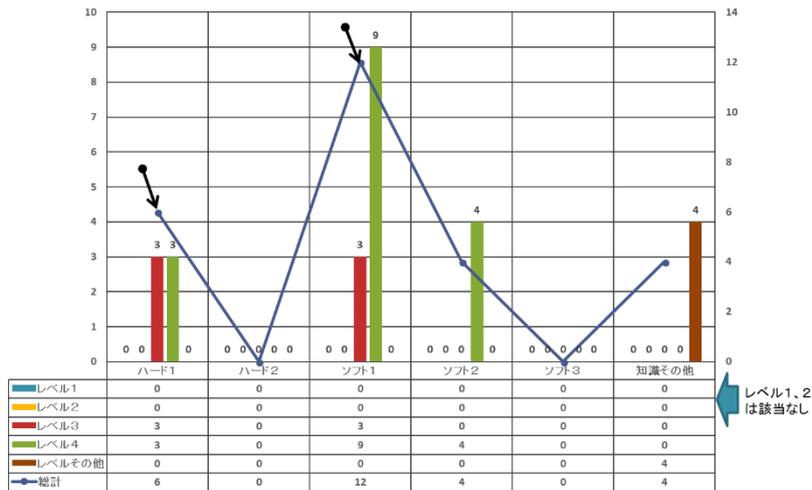
① 日本経済新聞



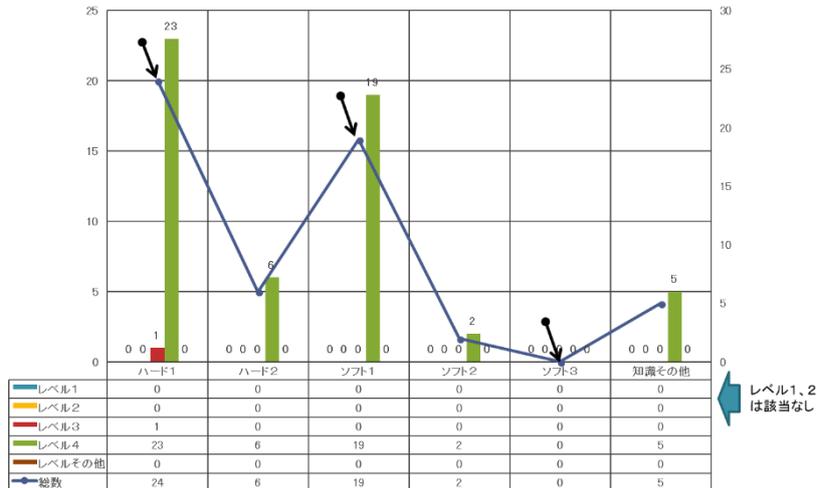
② 人工知能学会誌



③ Preferred Networks 社のホームページに開示された事例



④ 政府の取り組み事例 (NEDO)

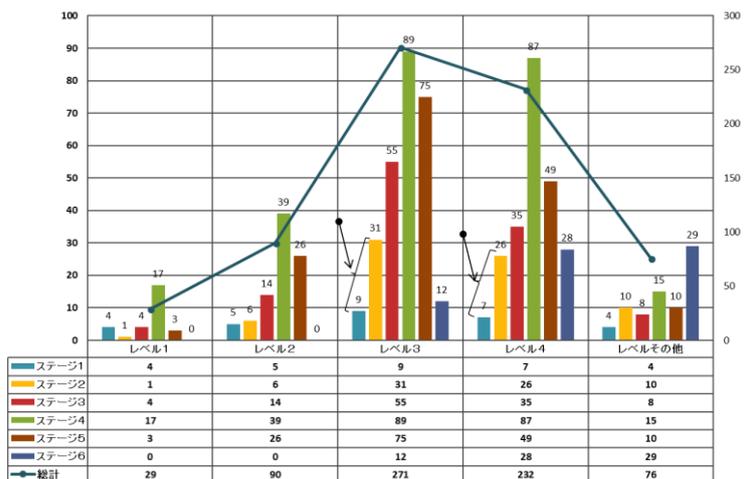


(2) 「発展段階レベル」と「研究開発ステージ」

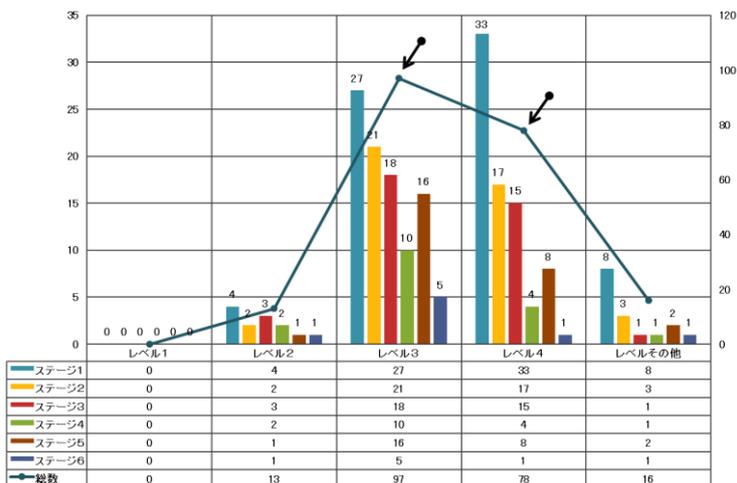
「発展段階レベル」と「研究開発ステージ」の関係を見ると、① 日本経済新聞、③ Preferred Networks 社の事例は、ステージ4以降が多く、既に製品化やサービス化が図られたものが多いことがわかる。ただし、① 日本経済新聞では、レベル3・4においては、ステージ2の事例も登場している。

また、② 人工知能学会誌、④ 政府の取り組み事例 (NEDO) ともステージ1ないし2の事例数が他のステージに比べ多く登場することから、これから新たに登場する事例などを探索するのであれば、これらの局面を深掘することが良いのではないかと考えられる。ただし、前述のとおり、① 日本経済新聞にも、ステージ2が登場することから、それらも合わせた探索が効果的でないかと考えられる。

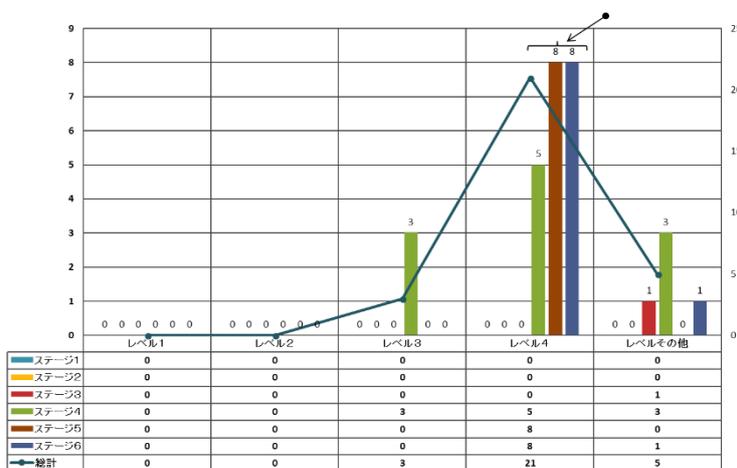
① 日本経済新聞



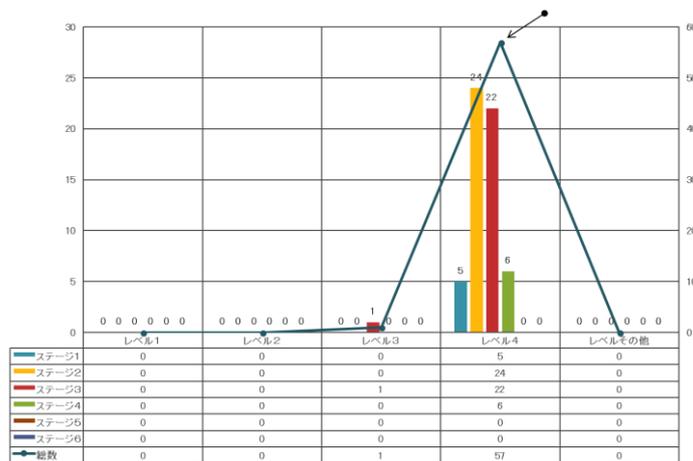
② 人工知能学会誌



③ Preferred Networks 社のホームページに開示された事例



④ 政府の取り組み事例 (NEDO)



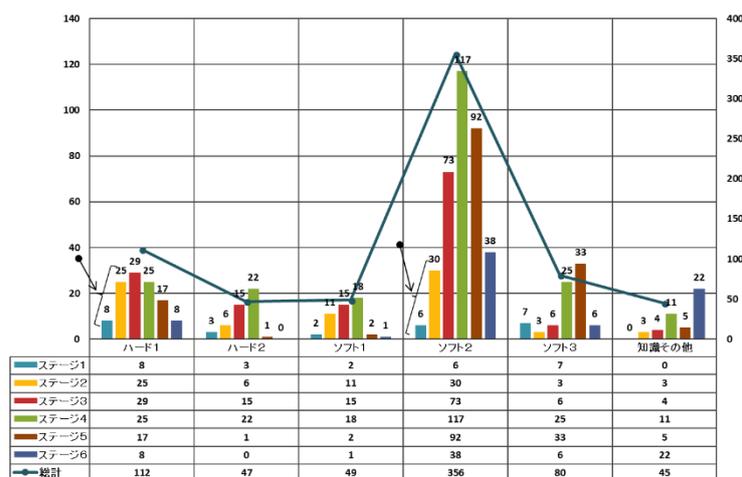
(3) 「知識カテゴリ」と「研究開発ステージ」

「知識カテゴリ」と「研究開発ステージ」においては、③ Preferred Networks 社では、ステージ1・2は登場していない。

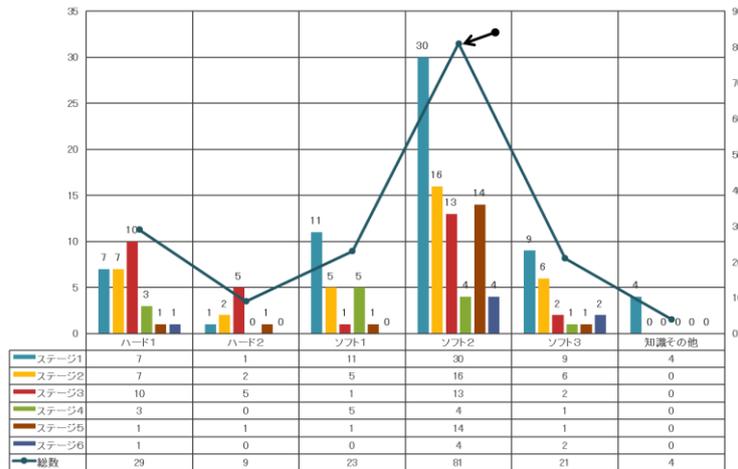
また、② 人工知能学会誌では、ソフト2のステージ1の事例が一番多く出現している。④ 政府の取り組み事例 (NEDO) においても、ステージ1が非常に多いが、その大半はハード1ないし2となっている。ちなみに、① 日本経済新聞では、ステージ1は非常に少ないが、ステージ2ではソフト2は比較的登場している。

「人間・社会システム」のステージ1・2に着目した場合、② 人工知能学会誌を深掘することが望ましいと考えられる。ただし① 日本経済新聞にもステージ1・2の人間・社会システムに関する事例も登場するため、合わせてそれらの探索も必要であると考えられる。

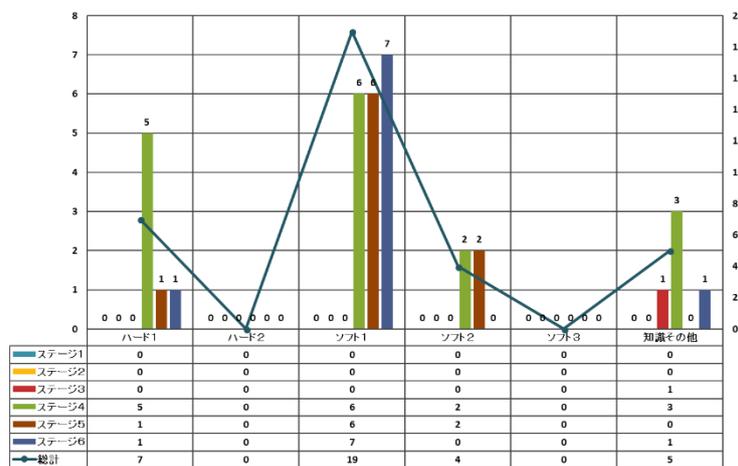
① 日本経済新聞



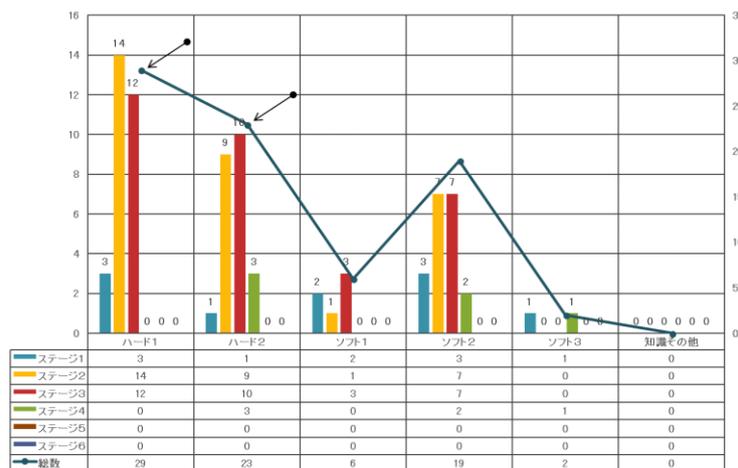
② 人工知能学会誌



③ Preferred Networks 社のホームページに開示された事例



④ 政府の取り組み事例 (NEDO)



3.1.4. その他

本分析において、用意した3つの分類軸毎に記事カテゴリを見ると、記事カテゴリにより登場する

各分類の項目（例えば、「発展段階レベル」のレベル4など）の比率の多いものには、特徴がみられる。

（1）発展段階レベル

- レベル1：ほとんどの記事カテゴリはこの項は登場しない。ただし、No.17 農業・漁業の比率は比較的高い。
- レベル2：比率の高い記事カテゴリはない（最大でも28%程度 No.14 インフラ検査・点検）。
- レベル3：どの記事カテゴリにもこの項は登場するが、比率の多い記事カテゴリとそうでない記事カテゴリに差がみられる。特に比率の高い記事カテゴリは、No.12 医療・創薬・ヘルスケア・介護、No.20 その他分野、No.5 ロボットの順になっている。
- レベル4：どの記事カテゴリにもこの項は登場するが、比率の多い記事カテゴリとそうでない記事カテゴリに差がみられる。特に比率の高い記事カテゴリは、No.1 AI スピーカー・会話・音声、No. 3 囲碁・将棋、No.2 自動運転の順になっている。
- その他：この項で登場する記事カテゴリは少ない。ただし、比率の高い記事カテゴリとしては、No.7 AI 人材（64%）となっている。

（2）知識カテゴリ

- ハード1：登場する記事カテゴリは少ない。また、全く登場しない記事カテゴリも多数存在する。ただし、登場する記事カテゴリは本項の比率が非常に高いものとなっている。なお、No.9 量子コンピュータ・スパコンは100%本項に該当する。また、No.14 インフラ検査・点検は9割弱が本項に該当する。
- ハード2：比率の高い記事カテゴリはない（最大でも25%程度、No.5 ロボット、No.17 農業・漁業）。さらに、全く登場しない記事カテゴリも多数存在する。
- ソフト1：5割を超える記事カテゴリはない（最大はNo.13 物流・配送・配車の40%）。また、全く登場しない記事カテゴリも多数存在する。
- ソフト2：登場する記事カテゴリが多いのがこの項である。ほぼ9割以上がこの項に該当する記事を有する記事カテゴリは6つ（No.1 AI スピーカー・会話・音声、No.2 自動運転、No.10 自動翻訳・通訳、No.11 コールセンター、No.12 医療・創薬・ヘルスケア、No.20 その他分野）存在する。
- ソフト3：登場する記事カテゴリは少ない。また、全く登場しない記事カテゴリも多数存在する。唯一、No.4 フィンテックの記事が50%以上本項に該当する。

（3）研究開発ステージ

ステージ全体では、基礎的な研究に相当するステージ1、2はどの記事カテゴリでも登場する割合は少ない。ただし、No.15 AI 用半導体は57%がステージ2である。

また、ステージ3の開発研究の割合が多い記事カテゴリとしては、No.2 自動運転（62%）、No.3 囲碁・将棋（45%）が存在する。その他の記事カテゴリは、ステージ4・5が中心となる。

ただし、ステージの判別ができない記事も多い記事カテゴリも存在する（No.1 AI スピーカー（53%）、No.5 ロボット、No.7 AI 人材（56%）、No.21 技術（46%）

表 13 新聞記事における各カテゴリ毎の3軸内の登場割合

No.	記事カテゴリ名	記事数	事案件数		人工知能のレベル					対象の知識カテゴリー						ステージ						
			対象	対象外	レベル1	レベル2	レベル3	レベル4	その他	ハード1	ハード2	ソフト1	ソフト2	ソフト3	その他	1	2	3	4	5	6	不明
1	AIスピーカー・会話・音声	67	90	26	0.063	0.13	0.795	0	0	0	0.954	0	0	0	0	0	0	0	0	0	0.53	
2	自動運転	32	32	0	0	0	0.7	0.3	0	0	0.871	0	0	0.128	0.08	0	0.82	0	0	0.05	0.14	
3	囲碁・将棋	20	18	2	0	0	0.26	0.738	0	0	0.176	0.235	0.588	0	0	0	0.45	0	0	0.3	0.1	
4	フィンテック(金融系含む)	53	53	0	0.06	0.195	0.421	0.322	0	0.04	0.44	0.515	0	0	0.06	0	0.23	0.23	0.6	0	0.1	
5	ロボット	31	13	18	0.143	0.849	0.21	0	0.076	0.25	0.061	0.45	0	0	0	0	0.23	0	0	0	0.58	
6	働き方・業務効率化・業務支援	103	107	5	0.171	0.509	0.221	0.077	0.134	0.088	0.143	0.45	0	0	0	0	0.21	0.38	0.22	0	0.1	
7	AI人材	27	22	5	0	0.14	0.21	0.643	0	0.063	0	0.31	0.62	0.563	0	0	0	0.15	0.26	0.58	0	
8	ドローン	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
9	量子コンピュータ・スパコン	11	10	1	0.111	0.22	0.667	0	1	0	0	0	0	0	0.13	0.19	0.19	0.13	0.13	0.13	0.1	
10	自動翻訳・通訳	7	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0.22	0.56	0.22	0	0	
11	コールセンター、自動応答、チャットボット	16	16	0	0.158	0.529	0.157	0.155	0	0	0	0	0	0	0	0	0.22	0.28	0.44	0	0	
12	医療・創薬・ヘルスケア・介護	48	44	4	0	0.719	0.281	0	0	0	0	0	0	0	0	0.28	0	0	0.2	0.24	0.1	
13	物流・配送・配車	13	14	0	0	0.65	0.35	0	0	0	0.4	0.4	0.1	0	0	0.11	0.39	0.44	0	0	0	
14	インフラ検査・点検	37	36	1	0.278	0.611	0.111	0	0.872	0.128	0	0	0	0	0.04	0.28	0.25	0.19	0.13	0.09	0	
15	AI用半導体	9	5	4	0	0	0.6	0	0	0	0.286	0.28	0.286	0	0.14	0.57	0.14	0.14	0	0	0	
16	スマート工場・自動工場	14	15	0	0.063	0.31	0.563	0.063	0.353	0.176	0.063	0.35	0	0.05	0	0	0.27	0.53	0.2	0	0	
17	農業・漁業	9	15	0	0.462	0.154	0.29	0.154	0.583	0.25	0	0	0	0	0	0	0.2	0.47	0	0	0.23	
18	防犯・防災	8	8	0	0.063	0	0.33	0.606	0	0	0.1	0.2	0	0	0	0	0.25	0.63	0	0	0.1	
19	生活	24	20	5	0.208	0.45	0.15	0.187	0	0.043	0.333	0.45	0.14	0	0	0.05	0.12	0.36	0.05	0.32	0	
20	その他分野	23	24	0	0	0.867	0.133	0	0	0	0.849	0	0	0	0.08	0.25	0.28	0.28	0.28	0.28	0.1	
21	技術	53	39	14	0.143	0.32	0.281	0.25	0	0.075	0.208	0.37	0.07	0.208	0	0	0.18	0.18	0.18	0.07	0.46	
22	トピック	196	186	23	0.102	0.22	0.25	0.347	0	0.09	0.32	0.12	0.281	0	0	0	0.12	0.12	0.17	0.21	0.21	
	合計	802	776	109																		

人工知能のレベル4の比率が多い記事カテゴリはNo. 1~3, 9, 15, 16

対象の知識カテゴリーソフト3の比率が多い記事カテゴリはNo. 4

(4) 記事の知識カテゴリ別

新聞記事の評価結果から、記事の知識カテゴリ毎に各軸の中で登場する記事の割合(0~1)で整理したところ、下表のような関係を見て取ることができた。近年人工知能関連で多くの話題提供を行っている、No. 1~3については、レベル4かつソフト2である記事が大半であった。また、人工知能レベル3かつソフト2である記事(下表のL3・S2)は主に企業で用いられるサービスや製品に関連するものであった。また、ソフト3を多く含む記事カテゴリとしては、No. 4 フィンテックのみが該当した。さらに、知識カテゴリでハード1を多く含むものは、基盤技術や社会インフラの領域に属する記事カテゴリとなっている。また、記事全般を通してステージ1・2の基礎的な研究要素に該当する記事は少ないことが見て取れる。

類似グループ	No.	記事カテゴリ名	人工知能のレベル					対象の知識のカテゴリー						ステージ						
			L1	L2	L3	L4	他	H1	H2	S1	S2	S3	他	ST1	ST2	ST3	ST4	ST5	ST6	他
L4・S2	1	AIスピーカー・会話・音声	0	0.07	0.14	0.8	0	0	0.02	0.02	0.98	0	0	0	0	0.05	0.4	0.01	0.01	0.53
	2	自動運転	0	0	0	0.7	0.3	0	0	0	0.88	0	0.13	0.03	0.05	0.62	0.05	0.05	0.05	0.14
	3	囲碁・将棋	0	0	0.26	0.74	0	0	0.18	0.24	0.59	0	0	0	0.05	0.45	0.05	0.3	0	0.15
	15	AI用半導体	0	0	0.4	0.6	0	0.14	0	0.29	0.29	0.29	0	0.14	0.57	0.14	0.14	0	0	0
	16	スマート工場・自動工場	0	0.06	0.31	0.56	0.06	0.35	0.18	0.06	0.35	0	0.06	0	0	0.27	0.53	0.2	0	0
L3/4・S3	4	フィンテック(金融系含む)	0.06	0.2	0.43	0.32	0	0	0.04	0	0.44	0.51	0	0.06	0.24	0.02	0.23	0.6	0.02	0.04
	10	自動翻訳・通訳	0	0.1	0.5	0.4	0	0	0	0	1	0	0	0	0	0	0.22	0.56	0.22	0
L3・S2	11	コールセンター、自動応答、チャットボット	0	0.16	0.53	0.16	0.16	0	0	0	1	0	0	0	0	0.22	0.28	0.44	0	0.06
	12	医療・創薬・ヘルスケア・介護	0	0	0.72	0.28	0	0	0	0	1	0	0	0	0.29	0.08	0.02	0.2	0.24	0.18
	20	その他分野	0	0	0.67	0.33	0	0.11	0	0	0.84	0	0.05	0.08	0.05	0.25	0.29	0	0.29	0
	18	防犯・防災	0.08	0	0.5	0.33	0.08	0.1	0.1	0.2	0.5	0	0.1	0	0	0.25	0.63	0	0	0.13
	5	ロボット	0	0.14	0.64	0.21	0	0.21	0.25	0.08	0.46	0	0	0	0.1	0.23	0.1	0	0	0.58
	6	働き方・業務効率化・業務支援	0.02	0.17	0.5	0.23	0.08	0.13	0.09	0.14	0.48	0.08	0.07	0.02	0.04	0.21	0.38	0.22	0.01	0.12
	19	生活	0	0.21	0.46	0.17	0.17	0	0.05	0.33	0.43	0.14	0.05	0	0.03	0.12	0.36	0.04	0.08	0.32
	13	物流・配送・配車	0.1	0.05	0.55	0.25	0.05	0	0	0.4	0.45	0.15	0	0	0.11	0.39	0.44	0.06	0	0
	14	インフラ検査・点検	0	0.28	0.61	0.11	0	0.38	0.13	0	0	0	0	0.04	0.28	0.28	0.19	0.11	0.09	0.02
	9	量子コンピュータ・スパコン	0	0.11	0.22	0.67	0	1	0	0	0	0	0	0.13	0.19	0.13	0.13	0.13	0.13	0.13
17	農業・漁業	0.46	0.15	0.23	0	0.15	0.58	0.25	0	0	0	0.17	0	0	0.2	0.47	0	0	0.33	
	7	AI人材	0	0	0.14	0.21	0.64	0	0.06	0	0.31	0.06	0.58	0	0	0.04	0.15	0.26	0.58	0
	8	ドローン	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	21	技術	0	0.14	0.32	0.29	0.25	0.06	0.08	0.21	0.38	0.08	0.21	0	0.07	0.16	0.19	0.05	0.07	0.46
	22	トピック	0.07	0.1	0.23	0.26	0.35	0.06	0.04	0.05	0.39	0.18	0.28	0.03	0.08	0.07	0.18	0.14	0.17	0.32

基礎研究的なステージにはない?(AI用半導体を除く)

■ : 登場割合が当該軸中50%以上

■ : 登場割合が当該軸中40%以上50%未満

3.2. 現状取り扱われている人工知能の実態について

実態分析を通して見えてきた人工知能は、巷で話題となっているような人類を凌駕するような進化を目指した人工知能だけではない。既存の機械の延長線上にあるような、より便利で効率的な社会を実現するために作られる人工知能など様々なものがみられた。

また、取り扱われている人工知能のレベルについても、第3世代型人工知能の中核をなすレベル4だけではなく、レベル1、2といった第2世代以前の人工知能も事例では取り扱われていた。

例えば、「3.1.2 事例パターン」から、「発展段階レベル」と「知識カテゴリ」の組み合わせが一意に各軸の項目が割りあてられているもの（レベルとカテゴリ双方が一意に割り振られているもの）のみを取り出し、対象の「知識カテゴリ」の各分類と人工知能の「発展段階レベル」に分けて整理すると下図のとおりとなる。なお、4つのデータソース全体での事例数は1163件であり、そのうち対象外は62件、すべて不明が316件、一意にレベルとカテゴリが割り振られたものは400件となっている。以降の考察は、この400件について詳細に行った。

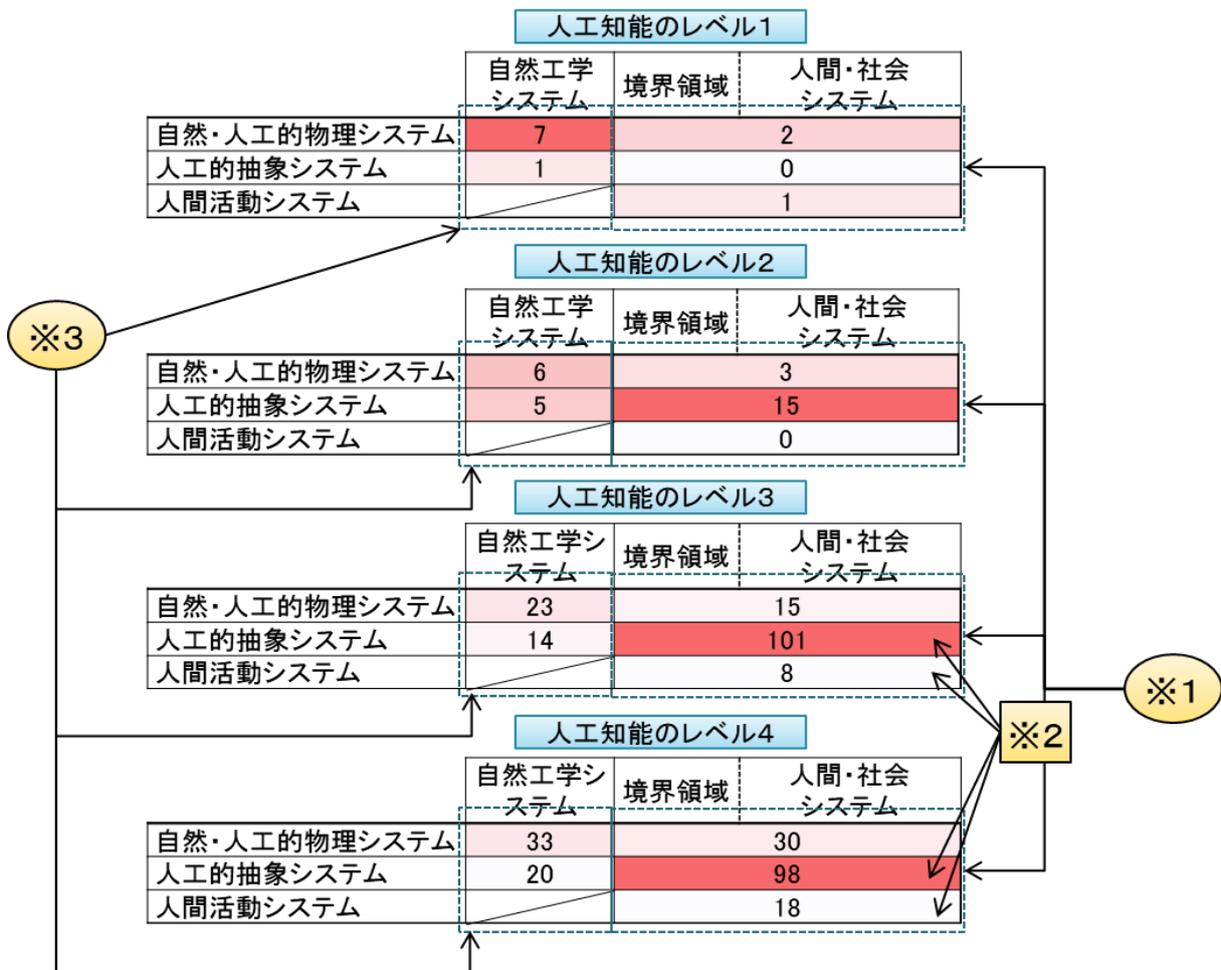


図4 取り上げられた事例のパターン（人工知能の発展段階レベル別）

人工知能の「発展段階レベル」が1、2では、ほとんどが自然・人工的物理システムを対象としたものである。

人間・社会システムを対象とした人工知能を実現するためには人工知能の「発展段階レベル」3、4は最低限必要となっていると考えられる。(図中※1)

ただし、その多くは人工的抽象システムであり、人間活動システムを対象としたものは少ない。(図中※2)

また、自然・工学システムにおいても、人工知能の「発展段階レベル」1、2に比べ、レベル3、4に該当する事例が比較的多く存在することから、「発展段階レベル」が高まったとしても自然工学システムへの適用は拡大しているものと考えられる。(図中※3)

※人工知能の「発展段階レベル」1において、人間活動システムが1件登場するがこれは事例内容の誤りだと考えられる。(外国の銀行におけるAIを活用した個人向け取引の学習に関する新聞記事で、対象のキーワードは「銀行口座管理」としてピックアップされた事例)

3.2.1. 自然システム・人工的物理システムの状況（ハード1）

本領域はフィジカルシステムに該当する領域である。本領域では、人工知能のレベルが1もしくは2に限定される事例の件数は多くないものの登場している。実際に取り上げられた事例において人工知能のレベル1や2では、自然環境情報の分析（環境データ、(牛の)生体データ、(牛の)行動記録、生育データ、生産時の温度管理、水田管理環境データ、温室環境データ、建設現場の風速・水位・温度などのデータなど）による建設現場（建機）、工場や農業への人工知能の適用といったキーワードがあげられる。なお、これらの事例のうち人工知能のレベル1はステージ4の製品開発が大半であり既に実用化された内容が中心であると考えられる。また、レベル2ではステージ3の研究開発段階が大半であり、レベル2であっても研究開発の途上にあるものも存在すると考えられる。

さらに、人工知能のレベル3、4においても多数の事例が登場している。事例で取り扱っている対象は、動植物情報の分析（植物の糖度だけでなく比重や色、醸造に関するデータ）、自然環境情報の分析（天候や曜日、発電量の過去のデータ、画像、色彩情報、元素情報、代謝情報、気象情報、土壤のプロファイル情報）、作業や環境情報の分析（熟練作業員が実際に操作する動きなどの情報、大量の交通データ）などのキーワードがあげられる。なお、これらの事例のステージは、人工知能のレベル3ではステージ1の科学研究、ステージ2の技術研究・開発も多くみられる。この傾向は、人工知能のレベル4でも同様である。これらの特徴としては、従来は取得することが困難であった大規模なデータ（ビッグデータ）やセンサーやデータ取得技術の進化（IoT）により様々なデータが取得可能になったことと、機械学習などの人工知能アルゴリズムの進化により、適用領域が広がった事例が多くみられる。事例の適用領域を見ると、工場や農業の生産現場、発電所や浄水場などの施設の管理、交通渋滞の制御など様々な領域に適用されている。さらに多くの問題解決を必要としており、人工知能の研究開発の発展が見込まれると考えられる。また、本領域は今後益々第3世代型人工知能の範囲で様々な研究開発が進むものと考えられる。

3.2.2. 自然システム・人工的物理システムの状況（ハード2）

本領域は主にフィジカルシステムに該当する領域ではあるが、実体概念による対象に人間・社会システムを含む領域である。本領域では、人工知能のレベル1、2の事例数はそれぞれ2件、3件となってお

り、登場数は少ない。実際に取り上げられた事例において人工知能のレベル1や2で取り扱っている対象のキーワードには、人の働き方に関する情報をもとにしたロボット化(生体情報、勤惰管理、健康管理、AI 管理人、銀行業務ベテラン知)といったものがみられる。これらは、人間・社会に関する領域ではあるが、業務の効率化といった観点であり、RPA* (ロボティック・プロセス・オートメーション) のような業務をソフトウェアロボットで置き換えるといったアプローチの基になる取り組みとみることができる。

さらに、人工知能のレベル3、4で取り上げられる対象のキーワードでは、流通や物流の機械化(無人店舗、スイカ決済、商品流通、在庫管理、マテリアルハンドリング(コンビニ、移動経路計画、把持対象物の選択、姿勢推定))、農業支援(環境データ、病害データ)、自動運転(自動走行、自動飛行)」、防犯関連の機械化(歩行者の属性、画像、行動する様子、安全・安心、生活現象モデリング(安全やヘルスケア))、ロボット(ロボット、きめの細かい動作、作業動作、不定形物操作、人間と相互理解、日常生活支援、人・機械協働生産)といった領域が見て取れる。また、これらのステージの多くは1～3であり、まだまだ研究開発段階であるとみることができる。本領域で登場する事例は、より高度でかつ社会生活に組み込まれるロボットなどの機械であり、今後ますます展開が期待される領域であると考えられる。

*RPAについて言及した事例は、本件調査では1件のみであり、本領域に分類されたものには登場していない。

3.2.3. 人工的抽象システムの状況(ソフト1)

本領域はサイバー・フィジカルシステムに該当する領域である。本領域では、人工知能のレベル1の事例数は1件となっており、ほとんど出現しない。また、人工知能のレベル2については、5件の事例が登場する。本項に登場する事例で取り扱っている対象のキーワードは、業務支援ソフト(事務処理、契約確認業務)、作業の解析結果の応用(物流センター)、ゲーム(サッカーの試合局面を離散的な状態遷移で表現するAIシステム、リアルタイムサッカー試合シミュレータ、高速で試合結果を計算するエンジン)、データ予測支援(微生物の吸収・代謝物の時系列データ)となっており、主に業務や何らかの行動のシミュレーション用いられるソフトウェアである。また、これらはステージ4、5の実用フェーズにある。

人工知能のレベル3、4においては自然システム・人工的物理システムに比べ数は少ないものの、14件、20件の事例が登場する。

登場する事例で取り扱っている対象のキーワードは、音声認識・検索、画像認識(物体・人物の認識、サイズ・形状認識、多形態画像対応)、業務支援ソフト(工業生産プロセス現象模擬、科学技術予測、建設施工計画、クレジットカード加盟店管理、物件相場、構造設計業務効率化)、新機能物質の開発、ゲーム(パラメータなどのデータの設計、キャラクターの推論・行動、キャラクターの学習、ゲームキャラクターの行動判断や身体パラメータの学習、自律的な最適パラメータの獲得、曲や歌詞の自動生成、解説文の自動生成、エージェントアーキテクチャ、仮想的な感覚器官により環境から得た情報をもとに行動を決定、プレイヤーのゲームスキル測定、定石も自分で発見、教えるデータがなくても学べる など)、作業支援ソフト(客による運行便検索システムの支援、訓練による体の使い方の最適化、演奏の習得、身体的学び、合唱や合奏・ピアノ演奏などの指導を支援するシステム、和太鼓のリズムを矯正するシステム、指導者による指摘、メタ認知による自己観察、アナロジーを組み込んだアブダクションシステム(この説明)、店舗器具修理指導 など)、脳のシミュレーション(全能ネットワーク分析、全能アーキテクチャ、記憶、ヒト海馬、脳内意味表現、社会知性)、AI用ソフトウェア(オープンソース、深層学習フレ

ームワーク、汎用配列計算ライブラリ、二階微分をサポートする関数、GPU 上での高速・省メモリな学習、GPU 向け行列計算ライブラリ、分散学習機能、マルチ GPU の集合通信ライブラリ、サンプルコード など) などより高度なシミュレーションをレベルの高い人工知能技術により実現する取り組みが多くみられる。これらの中には、人間の動作をシミュレートするだけではなく、デジタルゲームのような仮想空間においてのエージェントの行動を趣味レーションにより実現するといったサイバースステムのみを対象としたような事例がある。また、汎用的な計算ライブラリやオープンソースとしてソフトウェアを提供するような基盤構築の取り組みが見られる。さらに、人間の脳の仕組みを物理的にシミュレートする取り組みなど今後も継続して研究開発が行われると考えられる領域もみられる。ただし、これらの取り組みは人間の心に踏み込むというよりは、脳自体の仕組みの解明や人間動作のシミュレーションといった自然科学的な側面が現れている。また、ステージについても、AI 用ソフトウェアはステージ 5 (競争力開発)、6 (基盤研究・開発) が中心ではあるが、その他については、ステージ 1 (科学研究、主に脳のシミュレーション)、3 (開発研究) が多く、第 3 世代型人工知能の範囲での研究も進むと考えられるが、脳のシミュレーションなど更なる別の次元の人工知能として成果が期待される領域にあることが見て取れる。

3.2.4. 人工的抽象システムの状況 (ソフト 2)

本領域は境界領域と実体概念による対象に人間・社会システムを含む領域 (自然工学システムとの境界領域も含む) である。本領域においては、人工知能のレベル 1 の事例は登場せず、レベル 2 の事例も 15 件とレベル 3、4 に比較すると非常に少ない。事例件数の面から見ると、サイバー・フィジカルシステムが中心となるこれまでの領域 (ハード 1、ソフト 1) とは取り組み事例の傾向が異なるものと考えられる。また、全領域を通して、事例件数が一番多いのが本領域になる。これらから、本調査で取り上げた 4 つの局面において「現状の人工知能の適用領域としてはソフト 2 が注目されているのではないか?」という仮説が考えられる。

さらに本領域においてどのような対象や応用分野が見受けられるかを、登場する事例で取り扱っている対象のキーワードを中心に見ると以下のような結果になる。

(人工知能のレベル 2) 本領域に登場する事例で取り扱っている対象のキーワードには、次にあげる通り、人間に関連するサービスやソフトウェアを実現するものが多い。業務支援 (対人) (弁護士業務、スマートフォン (スマホ) で来店客を判別し、属性や購買履歴に応じ商品情報やクーポンを届ける、保険商品通販、着こなしの悩み相談、名刺管理、人脈づくり、無人カウンター、コールセンター、回答パターン、その他情報)、人事 (採用管理ツール、複数の面接官と応募者の予定、人材配置、退職可能性)、フィンテック (マンション購入と賃貸のどちらが得か判定)、業務支援 (定型書類業務、主要な経営指標 (売上高、運転資金の増減、想定支出額と手元資金のバランス、等)。これらには、ソフトウェアによる企業の業務支援系が多くみられるが、それらに共通してみられるのが、個々の人間を相手に対応をしなければならない業務や対応が求められるものになっている点である。

(人工知能のレベル 3) 本領域の人工知能のレベル 3 に該当する事例が一番多く登場する (101 件)。なお、他の領域においても登場していた対象のキーワードとしては、次のものがあげられる。音声認識 (AI

スピーカー含む）（音声解析、エコー、エコードット/プラス、音声認識言語処理、新型家電音声認識 など）、画像分析（細胞画像、がん発見、画像認識・医療診断、CT 検査、肺炎などの診断、内視鏡検査、病変の見落とし、MRI など画像データ、医療用画像、病理画像、動画投稿サイト、番組宣伝動画、動画認識、画像認識技術を使って出願された商標と似た商標がないかを調べる作業 など）、ロボット（日常生活支援、高齢者、障がい者 など）、自動運転。

これらは、対象となっている事例が直接人間に働きかけものであり、対象が自然工学的なものではなく、人間や社会そのものとしていることが見て取れる。

人工知能のレベル2にも登場していたものとしては、次のものがあげられる。レベル2以上にさらに個々人レベルへの働きかけが見られる。レベル2と異なるのは、大量のデータを用いた機械学習などによりこれらが、実現されている点にある。業務支援(対人)（コールセンター業務、広報・顧客サポート、不動産管理問合せチャット応答、対話型不動産売買の相談、コールセンター、チャットサービス、自動回答、問い合わせへの回答、店舗内店員適正配置、顧客との対話、最適な料金プラン等のソファでのタブレットによる表示、対話型eコマース、物件提案、一般からの特許に関する問い合わせへの回答案の作成、AI 裁判官）、人事・教育（医師国家試験、進路支援、人材採用、人事研修、能力評価、人材サービス、求職求人、心理分析、残業時間、有給休暇取得日数、上司が変わったか、自分で決める仕事の範囲 など）、フィンテック（AI 株式運用、為替動向予測、個人向け融資、与信業務や振り込み確認、取引先である40万社の支払い履歴や倒産情報、データ解析で信用力を判断、企業の業績や資金データ、寿命まで推定、事故現場の写真で保険金の支払額の見積もり、社会保険診療報酬、報酬支払い審査、取引先開拓、AI ファンド、投資銘柄選考、取引データの自動仕訳、AI による投資・運用助言、直感的な間取り検索 など）

他の領域や、レベル2に登場しないものとしては、以下のものがあげられる。これらの多くは社会生活に根差したサービスを提供することを目指したもので、さらにそれらを高度化するための基盤的な技術要素が含まれている。医療・介護・健康（医療診断支援サービス、要介護者ケアプラン・介護計画、医療・健康データ、健康指導システム、自治体のレセプト・診療報酬明細、疾患発症リスク、生活習慣病改善、糖尿病の合併症、糖尿病腎症の進行度合い予測、要介護者の状態判断、通院中断推定、狙い通りの病氣、薬剤投与、がん患者のゲノム解析、最適な治療法・薬 など）、自然言語処理（偽投稿、過激ツイート排除支援、自動翻訳、専門的文章、通訳、出願内容を技術分野ごとに分類、過去に同じような技術がないかの文献調査、特許出願された技術の内容を分析して分類、答弁書作成）、趣味・嗜好・感情・行動分析（人間の表情分析、対応履歴のテキストデータを読み込ませ顧客の喜びや不満が表れている部分を抽出、話しをしている人を見た人々が感じる印象の見極め、自社サイトへの来訪者の趣味・嗜好を分析、カード利用実績や居住地などの分析、お薦め情報の提示、自動的にお薦め物件の紹介、物件の「適正価格」提示、会員の購入履歴、生活習慣を分析、コト消費、リコメンドシステム、幸福感）、芸術（自動作曲）、防衛・治安（キラーロボット、LAWS、テロ対策、詐欺電話の解析・電話番号情報分析）。

（人工知能のレベル4）本領域の人工知能のレベル4に該当する事例はレベル3に次いで多く登場する（98件）。レベル3同様に、事例に登場していた対象のキーワードとして取り上げられたものは、ソフト1にて登場するものと同様のテーマ群（太字・下線）に該当するものがみられるが、ソフト1に登場して

いるものが、サイバー・フィジカルシステムのみであったが、本領域では人間や社会に係るデータが中心となっていることが分かる。例えば、特定の領域ではあるがゲームについては現実世界で人間と対決するレベルの人工知能を実現した DeepZenGo などの人工知能が含まれている。また、データ分析なども社会に生じている現象そのものを取り込み、さらにそれらの予測を行うシミュレーション系のソフトウェアが実現されている。

ゲーム (DeepZenGo、敵対的生成ネットワーク、アルファ碁、絶芸、人間の思考過程と競争)、データ予測支援 (景気予測、鉱工業生産指数の予測、消費統計、消費動向指数、備蓄量の変化、ショッピングセンターの売り上げの変化予測)、画像分析 (皮膚がんなど 14 種類の病気、MRI 画像、大腸がん、内視鏡画像、CT 画像、画像認識による不動産検索、約 5 3 3 万件の賃貸物件情報とそれにひも付く約 8 3 0 0 万点の画像データ)、業務支援 (対人) (多目的対話、先読み、接客・案内支援、教育・授業支援、選挙・大規模集会・会議等の合意形成支援、検索 (コールセンターオペレータ)、人事・教育 (転職者向けデータベース、マッチング、転職、採用)、自動運転 (自動運転、完全自動運転車)、音声認識 (AI スピーカー含む) (AI スピーカー、音声認識)

また、他の領域に登場しないものとしては、本領域の人工知能のレベル 3 と同じようなテーマ群が登場している。レベル 3 に登場するものとの違いは、認知や発想といった人間独自の領域のシミュレーションまで踏み込んだものがみられる点がある。

フィンテック (マネーフォワード、株式市場、トレーディング)、医療・介護・健康 (乳がん早期発見精度)、自然言語処理 (投稿チェック、自動翻訳・通訳、同時通訳、文間の意味関係、自然言語、契約書整理・論点抽出)、趣味・嗜好・感情・行動分析 (人の動きや趣向の傾向把握、ビデオカメラで撮影した観客の表情、満足度を数値化、テレビの視聴率を予測するシステム、地図情報と地域共通カードの顧客情報を組み合わせ分析、好みの店舗のクーポンや近くのイベント情報などを提供、事件・事故を未然に防ぐ予兆の発見、顧客ニーズに応えるヒット商品の開発、人やモノの移動の圧倒的効率化、匠の技・勘、需要予測などのシステム、顧客情報、推薦情報の選択)、芸術 (自動着色)、防衛・治安 (テロ対策、テロ動画の排除)

ソフト 2 の領域において登場している事例が現状の第 3 世代型においては大半を占めることが見て取れる。これらは、自然工学システムとの境界領域と人間・社会システムを対象としているケースではあるが、人間活動システムには踏み込んではいない。また、ステージについては 1～3 のみに該当するケースが半数以下 (62 件) であり、既に実用化されているステージ 4～が多いという特徴も有する。これらを鑑みると、「本領域は今後も事例は数多く登場するが、既に基礎的な研究からより応用面へシフトしているのではないか? (第 3 世代型の人工知能のゴール?)」と見ることができる。

3.2.5. 人間活動システムの状況 (ソフト 3)

本領域は境界領域と実体概念による対象に人間・社会システムを含む領域 (自然工学システムとの境界領域も含む) である。取り上げられた事例は日本経済新聞と人工知能学会誌のみからであり、Preferred Networks 社と NEDO の事例は登場しない。

また、本領域においては、人工知能のレベル1、2の事例は登場していない。（※図では人工知能のレベル1に1件の事例が含まれるが、誤りだと思われるため対象外とする。）レベル3、4においても、登場件数は他の領域に比べ少ない結果となっている。なお具体的には、レベル3が8件、レベル4が18件となっている。本項では、人間活動システムにどのような形で言及しているか、登場する事例の内容を個別に確認する。

人工知能のレベル3として登場する事例では、「チャットボット・音声認識」を対象のキーワードに含む事例では、「人間と人工知能の間に友情や恋愛感情が対話を楽しむうちに育まれる*1」といった内容である。この事例においては、人間の心の問題を取り上げているが、これらが絵空事ではない未来として語られてはいるレベルでしかない。また、人間の「不満・要望抽出」を対象のキーワードに含む事例は、人材関連企業の子会社が立ち上げた「不満買取センター」の事例*2であり、人間活動そのものではあるものの直接社会システムへ何かの働きかけを行うまでにはいっておらず、そこで得られたデータを販売するレベルにとどまっている。さらに、人工知能学会誌の「特集 広告とAI」では、人工知能の活用による新たな広告制作*3や広告への生成論的倫理の適用*4といった問題を取り上げているが特に後者は、議論と展望にあるように、思考実験の段階でかつ、ステージ6：基盤研究・開発のレベルにあるため、第3世代の人工知能の領域で実現できるかは微妙な状況にある。その他では、フィンテックの事例があげられるが、これらは不正送金や個人との与信管理など*5*6 個人の金融行動に関するものであり、人間活動システムに大きくかかわる内容となっている。ただし、本事例で踏み込んでいるレベルは、個人に関する個別の行動を対象にしてはいるが境界領域に大きく偏っておりコンピュータシステムとして実装することを念頭に実証実験を行っている段階にある。

*1 日本経済新聞社記事「気がつけばそこに（2）愛が生まれる日——絆、つなげるか（AIと世界）」

*2 日本経済新聞社記事「エン・ジャパン系、消費者の不満をAIで分析。」

*3 人工知能学会誌 2017 32 巻 4 号「人工知能による新しい広告クリエイティブ」

*4 人工知能学会誌 2017 32 巻 4 号「生成的倫理とその広告への適用について」

*5 日本経済新聞社記事「不正送金、AIで検知、住信SBIネット銀とNEC、地銀など向けシステム。」

*6 日本経済新聞社記事「三井住友カード、与信管理にAIを活用」

人工知能のレベル4として登場する事例から、「人間を支援する人工知能」、「人工知能と社会（人間）の関係」、「人工知能自身が思考する」、「全脳アーキテクチャ」、「フィンテック関連」などのテーマがあることが見て取れた。

「人間を支援する人工知能」については、新聞と人工知能学会誌からの得られた事例となっている。このテーマに含まれる内容については、他の領域においても業務支援ソフトウェアなどでも登場するが、本領域においては「異なる人種間の異文化ギャップを人工知能が支援する*7」といった、あたかも人間的なふるまいを目指した人工知能について述べられているものが登場する。しかし、これらについては取り組んでいる人が目指していることであり、第3世代型人工知能で実現できるかどうかについては言及されておらず、ステージも1：科学研究となっている。

*7 人工知能学会誌 2017 32 巻 1 号「「学習」支援と人工知能」

「人工知能と社会（人間）の関係」については、大半が人工知能学会誌からの事例となっている（新聞は1件のみ）。取り扱われている内容には、人工知能が作る社会の形*8 や AI と宗教という発展的課題*9、「人工知能と Emotion*10」、機械の時代にあっても「責任」が求められる重大な意思決定は人間が担い続けるか*11、「人間をうまく使うための人工知能技術が開発され、結果、人工知能が上位の存在となり人間を制御すること*11」、「AI やそれを取り入れたロボットは、人間の精神生活において、何か未曾有の異質さを持っているわけではない*12」、「患者側が AI を組み込んだロボットを望む場合も将来的には考えられる*12」などがあげられるが、多くは理想論や思考実験の範囲を超えておらず、現段階で人間活動システムに位置づけられる人工知能が第3世代型人工知能で実現できているとされている例は見取れない。新聞記事の事例においても、マイクロソフトの AI「テイ」が暴言を吐くような行動を起こした事件に対し、「例えば人種差別的なデータを学べば AI も人種差別的になる。*13」とマイクロソフト社副社長のコメントを取り上げているように、実際に動作している対話型の人工知能においても、人間の倫理観や社会観を実装することはできていないとみることができるとの発言がある。

また、AI が芸術を理解して人間にコンテンツを提供するという観点で、AI スピーカーの登場により、将来的には曲を選ぶのではなく、条件を決めるだけで新たな楽曲を作り出すことが可能になり、AI が音楽の楽しみ方そのものを変える世界について言及した記事*14 があるが、これらについても可能性の議論にとどまっている。

*8 人工知能学会誌 2017 32 巻 1 号「社会が先か、知能が先か」

*9 人工知能学会誌 2017 32 巻 1 号「人工知能と社会の接続点の設計：社会設計まで」

*10 人工知能学会誌 2017 32 巻 1 号「愛される人工知能を目指して」

*11：人工知能学会誌 2017 32 巻 1 号「人間を上手に使う人工知能」

*12：人工知能学会誌 2017 32 巻 5 号「異質な存在としての AI とその社会的受容」

*13 日本経済新聞社記事「AI 社会は信用できるか(Deep Insight)」

*14 日本経済新聞社記事「NIKKEI The STYLE-AI と生む究極の個の空間。」

「人工知能自身が思考する」については、新聞記事からの事例となっている。これらは、他の領域でも取り上げられたゲームに関する事例などであるが、それらの将来像へ踏み込んだ内容となっている。例えばグーグルが取り組んできた囲碁用 AI や「敵対的生成ネットワーク (GAN)」などによる深層学習の次の人工知能として、AI 同士が鍛え合ことで概念を自ら学ぶ力を得ることに言及している*16。

*16 日本経済新聞社記事「2つのAI、競い進化、深層学習の次は「敵対的生成NW」、自ら概念理解、実現も。」

「全脳アーキテクチャ」については、人工知能学会の取り組みとして紹介された事例であり、個別のタスクに特化した AI ではなく、汎用型の人工知能 (AGI) について議論がなされている。ここでは、特に AGI の領域において、「脳全体のアーキテクチャに学び人間のような汎用人工知能をつくる*17」ことが目標とされている。また、これらを実現するための提案として、「げっ歯類の海馬を参考にした深層学習の拡張モデルや前頭前野に関する知見を参考にしたモデルの提案*18」など第3世代型人工知の次の世代に向けた取り組みが議論されている。人間そのものの脳を実現することにより、出来上がる人工知能については、境界領域と人間・社会システムに係る内容として考えることができる。また、これらは取り組みはなされているものの、まだステージ1：科学研究～2：技術研究・開発段階である。

*17 人工知能学会誌 2017 32 巻 1 号「メゾスコピックコネクトームを A I アーキテクチャとして実装する」

*18 人工知能学会誌 2017 32 巻 2 号「全脳アーキテクチャ若手の会」

各テーマを見る限りにおいて本領域は、「人間活動システム」を含む系を対象としているが、特に人間・社会システム該当するもので取り上げられた事例の多くは、目標や目指す姿として描かれているものが多く、具体的に言及したものがないとみることができる。その理由として、第 3 世代型人工知能だけでこれらの具体イメージを描くことが困難なのではないかと考えられる。

4. レベル 4 人工知能の限界

本章では、これまでの鳥瞰的・包括的な事例分析をふまえて、それらの事例の中でも、「流通」、「金融」、「人事」の3機能領域に焦点をあて、各領域におけるレベル4人工知能の実装実態を深く掘り下げた分析を通じて、それぞれ、人工物理システム（ハード2類）、意思的人間を含む人工抽象システム（ソフト2類）、人間活動システム（ソフト3類）の3つの知識領域に関わるレベル4人工知能の限界の本質に迫っていきたい。

課題	主に関連する知識領域	事例	課題解決の見通し
制御性	人工物理システム（ハード2類）	流通・自動走行	レベル4での本質的解決は見通せていないものの、補完的な方法により、漸進的に実装可能な機能を拡大している
透明性	人工抽象システム（ソフト2類）	金融	レベル4での課題解決はソフト2類に関する限り見通せていない
創発性	人間活動システム（ソフト3類）	人事	レベル4での課題解決は見通せていない

4.1では、主に人工物理システム（ハード2類）に関わるレベル4人工知能の限界として、制御性の問題を取り上げる。人間が考えた機械的論理構造により組み立てるレベル1-3人工知能と異なり、レベル4人工知能は、プログラムが自ら、自律的に学習して振る舞うため、人間の開発者の意図による制御を維持するのが困難である。しかし、流通業界におけるレベル4の実装実態を通して、この制御性の問題が様々な方法によって補われ、対象機能を徐々に広げる現実的な対応が進んでいることを明らかにしていく。

4.2では、主に意思的人間を含んだ人工抽象システム（ソフト2類）に関わるレベル4人工知能の限界として、透明性の問題を取り上げる。透明性の問題は、記号接地問題や可読性の問題とも関わり、レベル4人工知能が、出力した結果の因果関係（理由）を説明できないブラックボックスとなっている問題である。第2章で取り上げた事例のうち金融市場に焦点をあて、トレーダーの人工知能による置換が急速に進行する金融市場で、実はレベル4人工知能がその中核的な部分を避けながら実装が進んでいる実態を明らかにし、そこからレベル4人工知能の人工抽象システム（ソフト2類）への適用の問題点についての一般論的示唆を見出す。

4.3では、主に人間活動システム（ソフト3類）に関わるレベル4人工知能の限界として、創発性の問題を取り上げる。第2章で取り上げた事例のうち人事に焦点をあて、レベル4人工知能が、人間のこころの奥深くに宿る差別意識・偏見を抽出しやすい特性をもつ一方、人間の社会性や直観的思考を抽出し

にくいという原理的特性をもつことを明らかにし、レベル4人工知能の人間活動システム(ソフト3類)への適用の問題について考える。

4.1 制御性の問題

4.1.1 スピードのロジックが拡散させるブラックボックス

世界最大の技術専門職組織を自認する米国電気電子技術者協会(The Institute of Electrical and Electronics Engineers: 以下 IEEE)は、人工知能分野における先進的な活動でも知られ、最近(2018年7月)もMITと共同で、人工知能のための新「物語」構築⁶を目的とした「拡張知能委員会⁷」を立ち上げている⁸。そのIEEEは、人工知能の倫理的課題の面でも世界でのイニシアティブを發揮しており、2016年に、世界のアカデミア、企業、行政から250人の専門家を集結し、「人工知能/自律システム倫理構想委員会⁹」を結成した¹⁰。2017年12月には、同委員会での1年間の議論を経て¹¹、パブリック・コメント向けペーパー『倫理的調和設計:人間ウェルビーイング中心の自律的知能システムのためのビジョン 第2版¹²』を起草し、レベル4人工知能の制御性に関して、次のように述べている¹³。

自律型システム(人工知能)の急成長を主導するディープラーニングは、ソフトウェアの世界を急速に「ブラックボックス」化させている。少なくとも予見可能な未来において、人工知能の開発者が、すべての状況下で人工知能を思い通りに振る舞わせるシステムを構築することは、おおよそ有り得そうにもない。それでもなお、人工知能のいかなる欠陥も障害も、あくまで、それを設計・構築・テスト・実装した人間が責任を負わなくてはならないのである。

レベル1-2人工知能は、開発者の論理的思考をIF-THEN文のロジック(Boolean logic)に落とし込んだプログラムで構成される。開発者は、何十人・何百人のチームで協働して統合的なシステムを構築し、プログラムの振る舞いを制御するための数十年間の経験に支えられた開発・テスト手法をもっている。それに対して、レベル4人工知能では、

- IF-THEN文ロジック(Boolean logic)は、蓋然性ロジック(Probabilistic logic)に代わり、
- 開発者の複合的な直観的・論理的思考は、コンピュータの単純なパターン分類マッピングに置き換えられ、
- 開発者の論理的・構造的設計は、サンプルデータに置き換えられている。

したがって、ここでは、原理的にいって、

⁶ “Build a New Narrative for AI”

⁷ The Council on Extended Intelligence : CXI

⁸ “MIT and IEEE Partner to Advance Extended Intelligence”, Lisa Morgan, Information Week, 7/17/2018

⁹ Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems

¹⁰ 『IEEEが「倫理的AI」の設計に関するレポートのドラフトを発表』 Stephanie Condon, ZDNet.com 翻訳:石橋啓一郎 2016年12月19日

¹¹ “IEEE launches ethical design guide for AI developers”, George Nott, Computerworld, 14 December, 2017

¹² ETHICALLY ALIGNED DESIGN :A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems V2.0, IEEE, <https://ethicsinaction.ieee.org/>

¹³ “ETHICALLY ALIGNED DESIGN : A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems V2.0”, IEEE, <https://ethicsinaction.ieee.org/>

- 真/偽 (truth / false) の明確な区別を喪失しており¹⁴、
- プログラムは概ねブラックボックスである。

そして、今日まだ私たちは、そのブラックボックスの振る舞いを有効に制御する手法をもっておらず、その開発途上にある¹⁵。

Google の研究開発ディレクターで、Google の AI シフトの中核的役割を担ってきたピーター・ノービックは、「我々は、なぜそんなものを使っているのか？」と自問し、「それは、シリコンバレーがスピードを求めるからだ」と 2017 年 5 月に開催された MIT Technology Review のフォーラム（『未来の知能システム構築のための最先端 AI ソリューション』）で語っている¹⁶。確かにレベル 4 人工知能は、特定の機能対象においては、人間の論理的思考を積み上げていくレベル 1-3 人工知能に比べて、何百倍も速く成果を出すことができる。しかし、その成果は、ブラックボックスが出力する結果の蓋然性が頼りで、人間の論理的制御がなかなか及ばない、正しいのか間違っているのか（真偽）の区別すら明確にできない。一步間違えれば、制御不能状態に陥って、大事故を引き起こしかねない危うさを内包しているのである。

4.1.2 「蓋然性による正確性」というチャレンジ

人工物理システム（ハード 2 類）に適用する人工知能は、「環境が仮定 X を充足するならば、振舞いは要件 Y を充足する」として要件定義される。レベル 1-3 人工知能においては、ひとつひとつの環境 X-要件 Y の組合せのすべては、既知かつ固定の機械モデルに関連して、言語、図式、チャート、スケッチ、グラフィック・アートなどによって表現される論理的構造に落とし込まれて定義されていた。それに対して、レベル 4 人工知能の場合、環境 X は、プログラムが「自律的」に自ら学習していく¹⁷。振舞い Y の正確性は、開発者がラベリングやアノテーション¹⁸加工した学習データによりプログラムを訓練する形をとって実装されるのであるが、対象とした環境におけるすべての環境仮定 X を網羅的に表現できる学習データを準備することは、人工物理システム（ハード 2 類）の場合、対象環境をよほど狭い範囲に限定しない限り、現実的には不可能なので、学習データは原理的に自ずとサンプルデータとなってしまう。サンプルデータによって表現される対象環境は不完全なものとなる。また、プログラムの学習は、蓋然性ロジックによって特徴量をマッピングする形でおこなわれるため、振舞い Y は原理的にいって蓋然的なものである。さらに、プログラムが、蓋然性ロジックによりニューロネットワーク上にマッピングしていく特徴量は、人間が中身をのぞきこんでもよくわからないため、不正な振る舞い Y が出力されても、そこに至るプロセスを検証したり、問題個所を特定して修正したりする方法が限定されている。つまり、どこまで対象環境の仮定 X を表現できているかがわからない不完全なデータに依存して、対象環境の網羅的な仮定

¹⁴ Peter Norvig, Director of Research for Google, "State-of-the-Art AI: Building Tomorrow's Intelligent Systems", MIT Technology Review, May 8, 2017

¹⁵ Peter Norvig, Director of Research for Google, "State-of-the-Art AI: Building Tomorrow's Intelligent Systems", MIT Technology Review, May 8, 2017

¹⁶ Peter Norvig, Director of Research for Google, "State-of-the-Art AI: Building Tomorrow's Intelligent Systems", MIT Technology Review, May 8, 2017

¹⁷ Stuart Russell, Daniel Dewey, Max Tegmark 「堅牢かつ有益な人工知能のための研究優先事項」 特集「AI 社会論」『人工知能』 VOL 32 NO.5 (2017 年 9 月号) 人工知能学会編

¹⁸ アノテーション：あるデータに対して関連する情報（メタデータ）を注釈として付与すること。AI 技術では、ディープラーニングを活用するための運用工程の 1 つであり、取得した大量のデータを識別および分類し、教師データ（正解データ）を作成する機能。NEDO Web サイトより：http://www.nedo.go.jp/news/press/AA5_100974.html

Xのうちの「一部」（どの一部かはわからない）をブラックボックス内で自律的に学習させ、その学習の蓋然的な結果として出力された振舞いYの「すべて」に対して、開発者は正確性を実装させなくてはならないという制御の難しさがあり、その意味での一種の原理的矛盾をはらんでいるのである。

そのうえ、レベル4人工知能における環境Xの学習には、大量のデータを必要とする。例えば、物体認識に広く用いられている畳み込みニューラル・ネットワーク（CNN）について、「ディープラーニング生みの親」ともよばれる¹⁹ジェフリー・ヒントン（トロント大学兼 Google）は、ディープラーニングは、データに関わる「指数関数的非効率性（exponential inefficiencies）」を備えていると指摘する²⁰。ヒントンがトロント大学ベクター研究所のセミナーで語ったところによると、空間移動に対する不変性（translational invariants）のために畳み込みニューラル・ネットワークに組み込んだ MAX プーリング層が、同時に大事な空間情報を失わせるよう作用しているため、空間の特徴が学習できない。そのため、オブジェクトの向き、場所が少しずつ異なる「指数関数的に」増量させたデータを必要とする構造になっているという²¹。もともと 1990 年代にはゆっくりとした試行的製品開発段階に入っていたニューラル・ネットワークが、2010 年代半ばから急速に華々しい成果をあげ始めたのは、ひとつには、ニューラル・ネットワークのこのデータ・ハングリーな性質を充足するだけの「ビッグデータ」がクラウド化の急速な進行によってサイバー空間上に蓄積されてきたという背景がある。しかし、2010 年代からのニューラル・ネットワーク台頭の中心舞台であった囲碁やチャットボット（意思的人間を含まない人工抽象システム（ソフト1類）と捉え得る）が、一定程度の蓋然性を達成していれば実用に耐えうるのに対して、人工物理システム（ハード2類）の場合、不正な振る舞いが人間の生命にも関わるため、正確性が要求され、そのために、レベル4プログラムの「指数関数的な」学習データ要求に対応しなくてはならない。現実には、学習データの準備が指数関数的な増量要求にデータの準備が追い付かず、不完全なデータで学習する結果となり、これは不正な振舞いYにつながり得る。

人工物理システム（ハード2類）への人工知能適用をめぐるのは、その倫理に関わる問題として、「トロッコ問題」などの議論がなされている。「ある人を助けるために他の人を犠牲にするのは許されるか？」といったもので、「急に道路上に5人が飛び出してきた場合、それを避けるためには歩道を歩いている人に突っ込まなければならない」などの形に置き換えられ多くの議論がなされている。しかし、たとえ、そうした倫理的な議論において一定の結論を出し、例えば、「常に死亡人数を最小化する振舞いを選択すること」といったような倫理原則を定めたとしても、人間ですら判断に迷う特殊な環境仮定Xの「すべて」を学習させるためのデータを網羅的に準備することは困難を極める。現実的には、定められた原理を学習させるに足りる十分な（指数関数的の増量された）学習データが準備することができず、これは不正な振る舞いYにつながり得る。最良の場合で、たとえ十分な学習データが準備できたとしても、原理を網羅

¹⁹ "Is AI Riding a One-Trick Pony?", James Somers, MIT Technology Review, September 29, 2017

²⁰ Sara Sabour, Nicholas Frosst, Geoffrey E. Hinton (Google Toronto) "Dynamic Routing Between Capsules", arXiv.org arXiv:1710.09829v2 [cs.CV] 7 Nov 2017

²¹ Geoffrey Hinton, "What is wrong with convolutional neural nets?" 2017-2018 Machine Learning Advances and Applications Seminar presented by the Vector Institute, University of Toronto, Canada, August 17, 2017, <http://www.fields.utoronto.ca/talks/What-wrong-convolutional-neural-nets>

的に完全に学習することができたかどうかを検証するすべが極めて限定されているため、結局は不正な振る舞い Y を招く可能性を残したままとなってしまう。

データ依存の問題は、「どれだけ」のデータを用意する必要があるかという量的な問題だけではなく、データの質的な問題、すなわち「どんな」データが必要なかが、(今のところ)よくわかっていないという問題でもある。量的問題は、アノテーションの手法開発や学習データを生産する人工知能の開発などによってある程度軽減できる見込みがあるのに対し、質的問題は、より深刻に制御性の問題にかかわる。例えば、画像に対し人間が気づかない程度の僅かな変更を加えるだけで、ニューラル・ネットワークによる画像認識を大きく誤動作させる(例えば、スクールバスをダチョウと認識させる²²)ことができる²³Adversarial Exampleはこの問題の一端を示している。Adversarial Exampleは、レベル4人工知能固有の問題とはいえないまでも、プログラムのレベル4人工知能の比率が高くなるほど、偽画像で「だましやすく」、脆弱になることが確かめられている²⁴。これを悪用して、現実世界の交通標識を誤認識させて事故を起こし、乗客に危害を加えるような攻撃をしかけるような現実的な攻撃行動ができるかどうかについては議論があるものの、攻撃対防御のコンテストも行われている²⁵。例えば、「ニューラル・ネットワークを Robust Physical Perturbations (RP2) にすると誤認識率を高めることができた」²⁶、「単純な分類器ではなく、物体検出によく使われる YOLO や Faster R-CNN といったモデルはだまされることなく” Stop ”の標識を認識することができた」²⁷、「adversarial example を直接生成するニューラル・ネットワークを作るというアプローチをとることにより攻撃性を高めることができた」²⁸などである。しかし、いずれの手法も、ニューラル・ネットワークの工夫によって、結果の表層的現象として、相対的により強い攻撃または防御能力を(たまたま)観察できたということを競い合っているのであって、このプロセスを通じて誤作動防止のための原理的な制御性が確保できつつあるわけではない。

Adversarial Example は、レベル4人工知能の質的なデータ依存性の問題の最も高度な一端を示したものにすぎず、高度ゆえに直接的に現実世界において悪用されて不正な振る舞いとして露出する可能性は低い。現実的には、この問題はもっと原初的なレベルで、例えば2018年3月のアリゾナにおけるUberの自動運転走行車の歩行者死亡事故のような形で露呈する。アリゾナの事故では、Uberの自動走行車は、自転車を押しながらビニールシートを羽織って歩いていたホームレスの女性を人間と検知できず、ビニールシートのような何かと誤認してしまったと『エコノミスト』誌が報じている²⁹。現実の道路では、ビニールシートや風船や凧や空の段ボール箱や様々な物体が舞うことはあり、ぬいぐるみや衣類などの物体

²² Anh Nguyen, Jason Yosinski, Jeff Clune, "Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images", arxiv.org, 2 Apr 2015

²³ 秋葉 拓哉 『NIPS'17 Adversarial Learning Competition に参戦しました』 Preferred Research 2018-04-11

²⁴ 野間口圭、黒米祐第、武田圭史、村井純 『FaceNet に対する Adversarial Example による意図的誤認識』 情報処理学会第79回全国大会 4W-01 セッション

²⁵2017年大会では、PFNは第4位に入賞した、秋葉 拓哉 『NIPS'17 Adversarial Learning Competition に参戦しました』 Preferred Research 2018-04-11

²⁶ Kevin Eykholt, Ivan Evtimov, Earlene Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, Dawn Song, "Robust Physical-World Attacks on Deep Learning Models", arXiv.org, 10 Apr 2018

²⁷ Jiajun Lu, Hussein Sibai, Evan Fabry, David Forsyth, "Standard detectors aren't (currently) fooled by physical adversarial stop signs", arXiv.org, 26 Oct 2017

²⁸ 秋葉 拓哉 『NIPS'17 Adversarial Learning Competition に参戦しました』 Preferred Research 2018-04-11

²⁹ "Uber's Fatal Crash Revealed a Self-Driving Blind Spot: Night Vision Heat-seeking sensors could make autonomous cars safer", Keith Naughton, The Economist, 2018年5月29日

が転がっていたり、落書きや落書きのような路面交通標識がペイントされていたり、あるいはダチョウや鹿その他の動物が横断したりすることもある。物体を検知するたびに停止していると追突事故を起こす危険な場合もあり、一方で、追突のリスクをとってでも優先して急停止しなくてはならない場合（例：電柱が倒れている）もあり、様々な環境が現実世界では起こり得る。しかし、そのまま走行しても問題ないクマのぬいぐるみやビニールシートのような物体と、人間（クマモンやスパイダーマンに変装している場合もある）や走行すると危険な物体（電柱やブロックや重そうな箱など）を確実に学習するために「どんな」データが必要なのかはよくわかっていない。そもそも、レベル 4 人工知能における物体識別は、Adversarial Example の事例にみられるように人間には視えないレベルでブラックボックスの中で自律的に学習されているのであるから、原理的にいって、それが近い将来に確実化・可視化でき、制御性の確保が見通せているとは言い難い。

データの質的な依存性の問題に関しては、Google の人工知能開発者チームも、モンテリオールで開催された『ニューラル情報処理システム会議 (NIPS) 2014』で、次のような懸念を表明している。「(ニューラル・ネットワークでは) 望ましい振舞いを外部データに依存せずにソフトウェアに組み込むことはできない。データの「癖」と抽象的な振舞いの不変性を切り離す手立てはほとんどなく、その結果(プログラム内での) 境界線の浸食が起こり、大きな技術的な負債を抱えこむことになる」³⁰。対象環境を表現するためのデータは、人工物理システムの場合、将来理論的に起こり得る事象も含めたとすると無限であるため、学習データが網羅的であることは現実的にありえず、学習データは、あくまで「サンプルデータ」でしかない。サンプルには、データの収集方法やそのデータが形成された場所の特性などによって、特定の「癖」が入り込む。レベル 4 人工知能は、その「癖」を特徴量に変換して、振舞い Y を出力することになるが、「癖」はニューラル・ネットワーク上のいたるところに分散して吸収されているから、特定の振る舞いを修正しようとしても、プログラム全体を作り直すことが必要になる。Google の同開発者チームは、この問題を「どこを変えても全部が変わる (Change Anything Change Everything : CACE)」問題と命名している³¹。これは、短期的には開発工程上の効率性の問題ではあるが、本質的なところで、サンプルデータに依存してブラックボックス内で自律的に学習するレベル 4 人工知能において制御性を確保することの難しさを示している。

また、ソフトウェアの「再利用性 (reproductivity)」の問題も指摘されている。振舞い Y が、データの「癖」に依存して出力され、データの「癖」と抽象的な振舞いの不変性が切り離されていないため、新しい方法で収集されたデータや、別の場所(物理的な場所とは限らない)で収集されたデータを用いると、元のデータで出力されていた振舞いを再現できなくなる³²という問題である。最新の深層強化学習により成果の得られたソースコード(乱数シード値、環境設定、ハイパーパラメータ³³、コード)をいくつかの

³⁰ D. Sculley et al (Google engineering team), "Machine learning: The high-interest credit card of technical debt", Neural Information Processing System (NIPS) 2014 Conference, SE4ML Workshop, DECEMBER 8 - 13, 2014

³¹ D. Sculley et al (Google engineering team), "Machine learning: The high-interest credit card of technical debt", Neural Information Processing System (NIPS) 2014 Conference, SE4ML Workshop, DECEMBER 8 - 13, 2014

³² Peter Norvig, Director of Research for Google, "State-of-the-Art AI: Building Tomorrow's Intelligent Systems", MIT Technology Review, May 8, 2017

³³ ハイパーパラメータ: ディープラーニングにおいて、人間が最初に設定しなければならない重みの初期値や学習率といったパラメータで、現状は試行錯誤で決定している。NEDO Web サイトより http://www.nedo.go.jp/news/press/AA5_100974.html

他の環境での再現性をテストしたモンテリオールの産学連携チーム（マギル大学とマイクロソフト）は、得られた結果から「深層強化学習の再利用性はほとんどないといってよい。とりわけ、標準的なベンチマーク環境における非決定性のテストにおいては、解釈すら難しい結果となった。」と結論付けている³⁴。CACE 問題とともに、この問題も、レベル 4 人工知能における正確性の保証や制御の難しさの一端を示している。

4.1.3 「想像をはるかに超える長い道のり」

トヨタ・リサーチ・インスティテュート（TRI）のギル・プラット最高経営責任者（CEO）は、「どんな環境でも AI が運転する完全自動運転の実用化は、スピード競争をすべきでない」と日経新聞のインタビューで語っている³⁵。また、お掃除ロボット『ルンバ』を開発した iRobot の創業者で、サブサンプレション・アーキテクチャの考案者でもあるロボット研究者のロドニー・アレン・ブルックスは、「ほとんどすべてのロボティクスと AI のイノベーションは、それがひろく社会実装されるには、開発現場の外にいる人達ばかりか、その中にいる人達の想像をさえ遥かに超えた、長い、長い道のりを歩まなくてはならない。実際の進捗は、苦痛を伴う、遅々としたものであり、AI は大量のデータを必要とするのである」³⁶と MIT Technology Review 誌に寄せたエッセイで述べている。

現実のレベル 4 人工知能の製品開発は、その長い道のりをゆっくりと着実に歩んでいるようである。第一に、対象環境を局所的な範囲に限定することにより、予測不能性をなんとか受容可能なレベルに抑える現実的な対応がなされている。例えば、DeNA の自動運転「ロボネコヤマト」の公道実証実験は、走行範囲を藤沢市の幅 500m、長さ 2km のごく狭い地域に限定し、ニトリの無人搬送ロボ（事例番号：769）は倉庫内に敷かれたレール上しか走行させないようにつくられている³⁷。プリファード・ネットワークスとファナックが共同開発した「ばら積みロボット」（事例番号：366）³⁸は、ロボットそのものが固定されていて、腕の動作範囲も 1m～3m 内に収められている³⁹など、そもそもの土台となる対象環境をまずごく狭い範囲に限定している。そのうえで、ハイパーパラメータやアノテーションの工夫、アルゴリズムの改良開発など精度を高めるための漸進的で地道な改善が進んでいる。「ばら積みロボット」（事例番号：366）の開発を指揮したファナックの稲葉清典ロボット事業本部長は、「正直なところ、まだ様々な課題があります。ばら積みロボットの機能を使用するためには、ある程度熟練した技術が必要です。例えば、どのワークから優先的に取り出すか、どのようにアプローチするかという判断を行うためのパラメータを調整する必要があります。現在、深層学習の技術にてそのような調整を行うことを取り組んでいます。故障予

³⁴ Peter Henderson, Riashat Islam, Philip Bachman, Joelle Pineau, Doina Precup, David Meger, “Deep Reinforcement Learning that Matters”, arXiv.org, 24 Nov 2017

³⁵ 「トヨタ、完全自動運転実用化には慎重、米研究所代表、AI 研究「事故死ゼロが目的」」 日本経済新聞（事例番号：408）

³⁶ Rodney Brooks, "The Seven Deadly Sins of AI Predictions", MIT Technology Review, November/December 2017 Issue

³⁷ 「ニトリの無人搬送ロボ、棚ごと移動し効率 5 倍、AI で出入庫を解析、今月稼働の自社施設に」 日本経済新聞（事例番号：769）

³⁸ 「製造ロボが深層学習、プリファード、ファナックと開発」 日本経済新聞（事例番号：366）の「ばら積みロボット」

³⁹ FANAC 『Robot i シリーズ General Catalogue』

知への応用にも取り組んでおり、壊れそうだという兆候の検出精度を更に上げる技術の開発も進めています。」とアジア・シームレス物流フォーラムで語っている⁴⁰。

第二に、データのアノテーションにも様々な手法開発が進んでいる。「ばら積みロボット」（事例番号:366）の場合、学習には数万枚の画像データを用いたが、データの学習方法を工夫することで、画像データに人間が詳細なアノテーションをしたのはそのうち数百枚程度に抑えることができたという⁴¹。第三に、アルゴリズムの開発も進んでいる。やはり「ばら積みロボット」（事例番号:366）の場合、物体が乱雑に置かれている環境が想定されていたため、物体が他の物体に隠れる様なオクルージョンが多く、布やビニールのような変形する物体も含まれており、古典的な画像処理手法では対処が難しい。そのため、ディープラーニングを用いた Semantic Segmentation という手法をベースにすることで、このような難しい環境下でも高精度に動作するシステムを実現している⁴²。また、認識アルゴリズムの開発に際しては、「バラ積みロボットが部品をつかめたら「成功」、取り損ねたら「失敗」という記録を動作ごとに残して、どういう場合にうまくつかめるかの教師データとし、これを深層学習させる。実際には、複数台のロボットが別々に動作し、複数台分の教師データを集めて深層学習させて認識アルゴリズムを開発している」という⁴³。

第四に、レベル4以外の人工知能とのハイブリッドで精度を高めているケースもある。例えば、スマートストアの実店舗営業と自動会計の実証実験を開始したトライアル・ホールディングスは、商品認識に、一般的に物体認識で最強と言われているレベル4人工知能だけではなく、レベル2-3のアルゴリズムも用いているという。同社の松下伸行執行役員によると、「ディープラーニングは、商品認識に適していない場合もある。例えば、お菓子の、「きのこの山」と「たけのこの里」のように似たパッケージは誤認識してしまう場合があり、この問題を、パノラマ写真を作る際に使う技術と同様の技術を使って、写真を分割しながら認識を繰り返しながら、一枚の写真に写る複数の商品の個数を数えることができる⁴⁴」ようにしたという。第五に、ハードウェアの機能改善への応用も進んでいる。「ばら積みロボット」（事例番号:366）では、熱変異補正を人工知能による熱変異予測によって実現している。熱変異補正は、工作機周囲温度や機械の動作中の発熱を温度センサーで検出し、温度変化による熱変位を学習技術を活用して予測し補正する機能で、温度センサーを用いない従来機能と比較して、加工精度が約40%改善し、また、温

⁴⁰ 『アジア・シームレス物流フォーラム 2017《ハイライト》 パネルディスカッション AIで物流ロボットがカンブリア爆発? 物流でのAI活用の可能性を探る ピッキング自動化・制御最適化で革命的進化へ』、月刊マテリアルフロー、2017年9月号、流通研究社

⁴¹ 米辻泰山（プレファード・ネットワークス / 東京大学）『CEATEC2016に出展してきました～ロボット編～』 Preferred Research 2016-10-20 <https://research.preferred.jp/2016/10/ceatec2016>に出展してきました～ロボット編～ /

⁴² 米辻泰山（プレファード・ネットワークス / 東京大学）『CEATEC2016に出展してきました～ロボット編～』 Preferred Research 2016-10-20 <https://research.preferred.jp/2016/10/ceatec2016>に出展してきました～ロボット編～ /

⁴³ 「PFNとファナック、バラ積みロボットのアルゴリズムを来年早々に市場へ投入 【特集】深層学習の雄、PFNの研究(1)」、多田 和司、日経 BigData、2016.12.07

⁴⁴ リテールAI研究会 『リアル店舗の逆襲 : 対アマゾンのAI戦略、人工知能(AI)が『流通・小売』を変える』 日経BP社 2018年06月

度センサーの配置や温度データの活用方法を工夫することで、万一温度センサーが故障した場合でも、加工を中断することなく最適な補正を継続することができるようになったという⁴⁵。

アルゴリズムの開発、アノテーションやハイパーパラメーター調整手法の開発、パーツとしてのレベル 2-3 との相互補完的利用など、人工知能の開発現場でのこうした地道な改善活動は、開発効率を着実に向上させ、それによってレベル 4 人工知能の社会受容可能な機能範囲を着実に拡大していくものと期待される。しかし、だからといって、その延長線上にレベル 4 人工知能の制御性の問題の根本的解決が見通せているわけではない。その延長線上にあるのは、あくまで適用対象（環境面と機能面）の漸進的拡張であり、世界中のどんな環境条件においても安全な稼働が制御される人工物理システムではない。自動走行を例にとるなら、その適用環境は、例えば、①走行可能な通行帯と走行方向が明示され固定的であること（発展途上国に行けば、この条件を満たさない通行帯はいたるところにある）、②道路標識や信号がねじ曲がったり、ひどく変色したりしていないこと、③交通法規に書かれていない暗黙のルールが存在しない（あっても開発段階で完全に把握されている）こと、④道路が陥没したり、洪水やがけ崩れや強風や水没や大地震や津波や火山噴火が起こっていないこと、⑤テロや暴動や戦争や内戦が勃発していないこと、⑥開発者が知らない・想像していない事象が起こらないこと、⑦開発者が学習データを準備しなかった事象が起こらないことなどの様々な条件を満たす「限定された環境」である。自動走行車が、世界中のいかなる地域のいかなる自然的・人工的・社会的条件のもとでも安全性が制御される「環境汎用」の人工物理システムに到達し、汎用完全自動走行車として世界中で販売可能になることは、レベル 4 人工知能の原理的特性を鑑みる限り、予見可能な将来においてあり得そうにもない。トヨタ・リサーチ・インスティテュート (TRI) のギル・プラット最高経営責任者 (CEO) は、同じ日本経済新聞のインタビューの中で、「トヨタなら悪いことが起きないという信頼性が価値になる」とも述べている⁴⁶。ビッグデータが集積し、金融市場のロジックに突き動かされるがゆえにスピードが求められるシリコンバレーで台頭してきたレベル 4 人工知能であるが、こと人工物理システム（ハード 2 類）の適用に関する限り、むしろそのスピードのロジックとは対極にある地道なモノづくり的改善の積み重ねのうえに、社会実装の着実な歩みが進んでいるようである。

4.2 透明性の問題

4.2.1 人工知能に置換されるトレーダー

2017 年 1 月 19 日、ハーバード大学科学センターのホール B では、『データ、マネー、アルゴリズム — 人工知能化する経済』と題する 2017 CSE Symposium が開催された⁴⁷。シンポジウムで最初の講演者を務めたゴールドマン・サックスの CFO 代行 (最高財務責任者代行) で前 CIO (最高情報責任者) のマーティ・チャベスは、金融市場における人工知能の実装に関して次のように述べている。「2000 年当時、ニューヨ

⁴⁵ GLOBIS 主催 『G1 サミット 2018』 第 5 分科会「AI が経営にもたらす創造と破壊～日本企業が勝つための戦略とは：パネルディスカッション～西川徹×松尾豊×村田大介×守安功×木村尚敬」 2018 年 2 月 11 日 <https://globis.jp/article/6314>

⁴⁶ 「トヨタ、完全自動運転実用化には慎重、米研究所代表、AI 研究「事故死ゼロが目的」」 日本経済新聞（事例番号:408）

⁴⁷ 2017 CSE Symposium <https://computefest.seas.harvard.edu/symposium-2017>

ーク本社では 600 人ものトレーダーが株式売買をしていていましたが、2017 年現在で本社に残っているトレーダーはわずか 2 人です。空いた席を埋めているのは、200 人のコンピューターエンジニアによって運用されている（人工知能による）自動トレーディング・プログラムです」⁴⁸。

金融市場でこのように人工知能の実装が先行的に進んでいるのは、他の業界と比べて、人工知能が実装しやすい技術的環境にあるためと考えられる。ソフト 2 類知識によって成立している金融市場は、人工知能にとってのインプットもアウトプットも共にサイバー空間であるため、ハード 1/2 類系製造業や流通業とは異なり、フィジカル空間からのセンシングなどによるインプットも、フィジカル空間への機械的なアクションのアウトプットも必要ない。そのため、データ収集や実験・試行のコストが製造業に比べてかなり低い。また、レベル 3/4 人工知能のための学習データとしては、世界中の金融機関、取引所、証券会社、個人投資家、事業法人、各国中央銀行、各国政府財務・統計部門などから毎秒ごとに大量に排出される膨大なデータ（為替/債権/株式/投資信託の発注・約定・発行データ、企業財務データ、アナリスト・レポート、ニュース、経済・産業統計など）が数十年にわたり、中立的なメディアによって集積されており、使いやすく構造化された状態で誰でも入手することができる。そのデータから、レベル 1-2 人工知能で簡単に算出できる株価やポートフォリオの相対的パフォーマンスが、そのままレベル 4 人工知能の「報酬」や「教師」として利用もできる。新たなセンシング技術を開発する必要はなく、新たに自ら学習データを収集・蓄積する必要性は相対的に低い。

そのうえ、同じくソフト 2 類の知識領域であっても、例えば広告業界の場合は、人工知能にフィードできる学習データが IT ビッグ 5 の寡占状態にあるが、金融業界の場合、ほとんど同じ量と粒度の学習データが誰にでも入手できるため、スタートアップやプラットフォームをもたない企業にとっても、広告業界ほどのハンディがない。ほぼ一線のスタートラインに立って、迅速に研究開発から製品化に取り組むことができ、それによって競争が促進され、人工知能の実装はさらに加速する。

本節では、このように先行的・急速に人工知能の実装が進む金融市場を事例として取り上げ、その深堀りの分析からレベル 4 人工知能の「意思的人間を含む人工抽象システム（ソフト 2 類）」への適用の問題を考えていく。なお、前章まででは、金融市場を FinTech に含めて論じていたが、FinTech には不動産が含まれており、土地・建物という物理的対象を人工抽象システム（ソフト 2 類）により取引する不動産と為替や株式といった抽象物を人工抽象システム（ソフト 2 類）で取引する金融市場とは性質が異なるため、本章では、分析の範囲を FinTech から「金融市場」に限定した。

4.2.2 人工知能に置換されないファンドマネージャー

ゴールドマン・サックスの 600 人のトレーダーを 200 人のコンピューターエンジニアに置換した金融市場であるが、この進んだレベル 4 人工知能の実装を、金融業界の機能分野別に分解すると、実装度合にかなりの偏りがある。ここでいう金融市場の機能分野とは、金融業界で広く一般的に用いられている市

⁴⁸ Nanette Byrnes “As Goldman Embraces Automation, Even the Masters of the Universe Are Threatened”, MIT Technology Review, February 7, 2017

場の機能セグメントで、以下の4機能（人間がその機能を担う場合は「職種」といってもよい）を意味する⁴⁹。

- ウェルス・マネージャー：個人投資家に対して資産運用のアドバイスを提供することを主機能とする。一部、一任されて資産運用をおこなう場合もある。日本ではファイナンシャル・アドバイザーとよばれることの方が多い。本調査では、この機能に融資、口座管理も含めている。この職種に求められる能力：顧客から信頼され、好かれる営業力、正確性など
- ファンドマネージャー：個人投資家や年金基金から委託をうけ資金を運用する機能。個人投資家向けの場合は、主に投資信託やミューチュアル・ファンドの商品形態をとる。この職種に求められる能力：企画力、分析力・判断力などのほか、（ほとんどの時間を投資家への説明やそのための資料作成に費やすため）公正さ、緻密さ、（投資家との）コミュニケーション能力
- トレーダー：ファンドマネージャーからの指示をうけ、また、金融機関の自己資金の運用を任せられ、株式・為替・債権などの金融商品の売買をおこなう。この職種に求められる能力：論理的思考による判断・行動が、ほとんど反射的にできるような瞬発力
- アナリスト：第一次的にはファンドマネージャーに対し、第二次的には個人投資家に対し、株価、為替、債権、経済に対する自身の（独自の）見立てを提供する。この職種に求められる能力：定量的分析力、定性的洞察力、文章表現力、（企業経営者や投資家との）コミュニケーション能力

第2章で概観した日本経済新聞の事例のうち、金融市場に関連する事例は34事例あった。この34事例を、4機能分野別のレベル4人工知能実装の度合は下記の表のとおりとなる。レベル4の実装が最も進んでいるのはトレーダー機能で、7件の事例すべてがレベル4であった。レベル4の実装が最も遅れているのはファンドマネージャー機能で4件の事例の1件しかレベル4が実装されていなかった。アナリスト機能は、8件の事例のうち37.5%がレベル4実装で、ウェルス・マネージャー機能は15件のうち、33.3%がレベル4実装であった。なお、4機能すべての事例において、ステージは、4（製品化ステージ）または5（競争力開発ステージ）であったため、ステージの違いは実装度合の考慮に入れていない。（ウェルス・マネージャー機能だけがステージの中間値が4となっているのは、個人が顧客であるウェルス・マネージャー機能においては、広報的目的で製品ラウンチ前にプレスリリースする傾向が強いためであり、実態的な実装ステージを反映したものではないと考えられる。）

⁴⁹ これ以外にインベストメント・バンカー機能を含めて合計5機能に分類する方法が一般的であるが、本調査ではインベストメント・バンカー事例が1件もなかったため、省略した。インベストメント・バンカーは、企業の資金調達、IPO、M&Aの仲介などをおこなう機能であり、一人当たり一日数十件の投資判断をおこなうトレーダーなどと違って、年間せいぜい数件の取引しかクローズしないため、人工知能化の経済性は低く、レベル4の実装は進んでいないものと考えられる

金融市場での機能	人工知能レベル（中間値）	ステージ（中間値）	事案件数	レベル 4 実装事例比率
トレーダー	4	5	7	100.0%
アナリスト	3	5	8	37.5%
ウェルス・マネージャー	3	4	15	33.3%
ファンドマネージャー	2	5	4	25.0%

人間のトレーダーが人工知能で急速に置換されている実態については前頁で述べたが、この機能分野別のレベル 4 人工知能実装度合の違いに（あたかも）相関するかのようには、人間のファンドマネージャーの人工知能による置換は進んでいない。ゴールドマン・サックスのニューヨーク本社の方のファンドマネージャー、オサマン・アリは、「他の業界では、人工知能が（職を奪うなどの）破壊的影響をもたらしているかもしれない。しかし、ことファンド・マネジメント業界に関して言えば、ビッグデータから新しいデータソースを見出すなど、我々はすでに機械学習と人工知能を（人間の）ファンド・マネジメント業務の中にうまく取り込み、最適な投資意思決定のために役立てている」とモーニングスターのインタビューに答え⁵⁰、人間のファンドマネージャーは存続したうえで、むしろ人工知能をうまく使う側に回っていることを示唆した。日本においても、ファンドマネージャーは、一般社団法人日本投資顧問協会によると、2017年3月末時点で2,140人⁵¹となっており、これは2016年3月末から11.62%増⁵²、2000年3月末の1,289人⁵³からは66%増と、むしろ著しい増加傾向にある。

なぜ、一方では、レベル 4 人工知能が人間の大部分を置換するほどに進行する金融業界においても、ファンドマネージャー機能ではレベル 4 実装が進まないのか。日経新聞のファンドマネージャー事例の中では、レベル 4 人工知能を使用している事例が1件、レベル 4 を他レベルと併用している可能性がある事例が1件あった。その詳細を調べてみると、2事例ともに、近年のバズ・ワード「人工知能」をマーケティング的に利用しつつも、ファンド組成・運用実体において全体のリターンとリスクに影響を及ぼさないよう、注意深く、試行的、限定的にレベル 4 を実装させようとする、担当ファンドマネージャーの配慮のあとがうかがえた。

⁵⁰ "Goldman Sachs: How Big Data, AI Benefit Investors", Morningstar, 19 Jun, 2018 (インタビューは Ben Johnson) <https://www.morningstar.com/videos/869561/goldman-sachs-how-big-data-ai-benefit-investors.html>

⁵¹ 一般投資一任業に限る、投資一任業でもラップ業務、不動産関連特定投資運用業は含まない

⁵² 一般社団法人 日本投資顧問協会 投資運用に係わる統計数値 http://www.jiaa.or.jp/profile/unyou_youran.html
<http://www.jiaa.or.jp/profile/pdf/youran29/hyousimokuji.pdf>

⁵³ 一般社団法人 日本投資顧問協会 投資一任業者の役職員状況 <http://www.jiaa.or.jp/toukei/index2d.html>
<http://www.jiaa.or.jp/toukei/pdf/itinin0918.pdf>

レベル 4 人工知能を明示的に使用している事例であったアセットマネジメント One の「AI(人工知能)活用型世界株ファンド」(事例番号:296⁵⁴)は、「ディープラーニング・モデル」をマーケティングの前面に露出し、企業財務指標(PER、配当利回り、売上高など)からレベル 4 人工知能による特徴量を抽出して、ポートフォリオ提案を出力しているが、人間のファンドマネージャーが同じデータを分析したうえで伝統的なポートフォリオ組成手法により最終のポートフォリオを組成している⁵⁵。なお、ファンドの 6 か月トータル・リターンは、ベンチマーク(トムソン・ロイター・グローバル・イクイティ・インデックス⁵⁶)を大きく下回っている。(ただし、レベル 4 人工知能を全く使用していなくても、もっとパフォーマンスの低いファンドはいくらでもあるので、このファンドのパフォーマンスの低さはレベル 4 の問題とは関係ない。)

レベル 4 を他レベルと併用している可能性がある事例であったゴールドマン・サックスの「GS グローバル・ビッグデータ投資戦略」(事例番号:606⁵⁷)は、もっと注意深い。ビッグデータ(リサーチ・レポート、決算説明会議事録、ウェブ・アクセス量、ニュース記事、アナリスト・レポートなど)から、「機械学習を通じた文章解析」によって、センチメント(市場心理)とモメンタム(勢い)を抽出している⁵⁸が、このアウトプットは、数万銘柄のユニバースの第一次銘柄フィルタリングに利用されているのみである。その第一次フィルタリングにおいても、収益力/バリュースコア、成長力/グロース指標や株価の勢い/モメンタム指標などの伝統的なモデルの補完としての利用にとどめられている。総合的にみて、ファンド組成・運用の大部分を伝統的なレベル 2/3 人工知能によりおこない、伝統的な手法では見落とされてきたアルファ(市場平均を上回る収益率)の発見のために試行的に⁵⁹レベル 3/4 人工知能を組み込んでいる。なお、ファンドの 1 年トータル・リターンは、ベンチマーク(トムソン・ロイター・グローバル・イクイティ・インデックス⁶⁰)とほぼ同レベルのパフォーマンスを達成している。

4.2.3 「責任者はどちらにいらっしゃいますか？」

レベル 4 人工知能実装に対するファンドマネージャーのこうした注意深さは、どこからくるのだろうか。このことを突き止めるために、ファンドマネージャー機能を全面的にレベル 4 人工知能に置換するための仮定的・簡易的な人工知能化システム企画分析を以下におこなうこととした。

最初に、現在ファンドマネージャーが主にレベル 1-3 人工知能を使用して実行しているワークフローを明らかにする必要がある。ファンドマネージャーの柱となるのは、銘柄選定タスクで、それは、次のような論理的プロセスをとる。

⁵⁴ 「AI 投信、安定収益で成長…分析対象で独自色競う、アセマネ One、深層学習を活用、大和投信、購買情報を分析。」日本経済新聞(事例番号:296)

⁵⁵ アセットマネジメント One『AI(人工知能)活用型世界株ファンド』「目論見書」及び「販売資料」

⁵⁶ トムソン・ロイター <http://online.thomsonreuters.com/indices/>

⁵⁷ 「注目投信担当に聞く(上)ゴールドマンサックス・アセット・マネジメント諏訪部貴郭嗣一運用に AI 活用」日本経済新聞社(事例番号:606)

⁵⁸ ゴールドマン・サックス『GS グローバル・ビッグデータ投資戦略』「目論見書」及び「商品詳細資料」

⁵⁹ 「常に 30 から 40 の新しい切り口を開拓するプロジェクトが走っており、年 2~3 回のペースで追加や改良を加えている」(担当ファンドマネージャーの諏訪部貴郭嗣氏)(事例番号:606、記事タイトル:「注目投信担当に聞く(上)ゴールドマンサックス・アセット・マネジメント諏訪部貴郭嗣一運用に AI 活用」)

⁶⁰ トムソン・ロイター <http://online.thomsonreuters.com/indices/>

- ① 最初に、投資方針を決定する。例えば、パフォーマンスに関連して「日経 225 連動（パッシブ運用とよんでいる）」「成長株」「割安銘柄」など、投資テーマに関連して「ESG」「人工知能の製品開発に積極的な企業」など、投資対象に関連して「アジア株」「新興国株」「G7 株」などである
 - ② 定められた投資方針を充足する指標を、何万種類もの既存のレベル 1-3 人工知能によって作成された指標（企業財務指標、経済指標、産業指標、株価指標など）から選定し、場合によっては新たな指標を自ら開発し、それらの指標に加重をかけ、総合して単一の論理構造をもった基準に集約し、その基準によって投資方針に適合する銘柄の候補を選定する
 - ③ 資本資産価格モデル(CAPM)などの既存のレベル 2-3 人工知能によって作成されたポートフォリオ理論を用いて、全体としてリスク・リターンが最適なポートフォリオを組成する
- といった具合である。①～③により定義されたポートフォリオ組成基準は、初期ファンド組成後、運用時の銘柄入れ替えにも基本的に適用される。ファンドマネージャーのワークフローの中でもうひとつの柱となるのは、ドキュメンテーションとコミュニケーションである。銘柄選定タスクにおいて組み立てられたファンドの仕様を投資家向け、経営者向け、監督官庁向けに説明するタスクで、次のような作業が含まれる。
- ④ ファンド組成・運用後の銘柄入れ替え基準が、経営者や監督官庁に説明可能なように、また、他のファンドマネージャーに業務を引き継げるよう、論理的な資料として詳細が記録する
 - ⑤ 同時に、投資家に対して、銘柄選定の基準とそれに関わるリスク（株価変動リスク、信用リスク、投資方針に関わるリスク（テーマ型ファンドの場合の変動率の高さなど）、カントリーリスク、流動性リスク、金利リスク、為替リスクなど）を、レベル 1-3 人工知能によって作成されたリスク指標にそって明示的に目論見書上、および、投資家や販売店向けとの PR やコミュニケーションの場において説明する

この説明タスクは、日本でも、北米でも、EU でも、世界中ほとんどの国の金融法規（日本の場合は金融証券取引法とその関連政令）で義務付けられている。ファンドマネージャーの勝手な個人的気分や好みでファンドが組成されてリスクがコントロールされていないファンドができたり、あるいは勝手に頻繁に銘柄入れ替えされて証券会社への手数料負担が重くのしかかってきたりして、投資家が明示的に許容していないリスクを負わされることのないよう、投資家保護の観点から法制化されている。

このような論理プロセスを経て、ファンドが組成され運用される結果、運用後のある時点で、もしファンドの価格がひどく下落して投資家からの問合せや苦情を受けた場合でも、それは「市場リスクによるもの、すなわち市場全体が同じくらいに下落したためです」とか、「急激に円高にふれたため（為替リスク）」であるとか、「人工知能バブルがはじけたため（テーマ型ファンに関わるリスク）」など、レベル 1-3 人工知能で作成された指標に基づいて説明することができる。投資家は、そのファンドに投資する時点で、事前に「目論見書記載のリスクを理解し許容したうえで投資します」との契約書に同意しているので、論理的にはそのひどい下落を許容するしかなく、苦情を申し立てる法的・論理的根拠をもたない。

それでは次に、このファンドマネージャーの銘柄選定タスクをレベル 4 人工知能で全面的に置換することを試みる。

- ① 最初の投資方針決定タスクのレベル 4 人工知能による置換を検討する。これは、マーケティング的なタスクでもあるので、マーケティングにレベル 4 人工知能を利用した他の業界の事例を参考にすることになる。例えば、小売業の店舗における顧客行動の映像データと POS の売上データをレベル 4 人工知能で分析し、マーケティングに活用する事例がある。しかし、ファンドの場合は、小売業とは違って、レベル 4 人工知能が得意な画像や音声などのビッグデータは見当たらない。日本市場での年間新規設定投資信託数は 500-700 本であり⁶¹、すべて論理化されたデータである。全国の店舗で毎秒何百人の来訪者があり、何百種類の商品が売れていく小売業に比べると、データの量としても質としてもレベル 4 人工知能には全く不向きである。したがって、投資方針決定タスクは、伝統的な消費者行動分析・マーケティング分析手法の方が有効な投資方針が策定しうると結論づけることとする。
- ② 次に投資方針を充足する指標の選定であるが、前段の投資方針決定タスクで投資対象市場が決定されているので、その市場のマーケット・データを入力データとし、ポートフォリオのパフォーマンス（リスク=リターンの最適化）を学習データとして、ニューラル・ネットワークを訓練する。そして、銘柄レベルとポートフォリオ・レベルの両方で、何度もバックテストを繰り返して、モデルと学習データの調整を行い、モデルのパフォーマンスを高める調整をおこなっていく。
- ③ を完了すれば（「完了」といっても、バックテストでそこそこのパフォーマンスが達成できたら切り上げないと、マーケティング部門から矢の催促がくる）、次のタスクである銘柄選定とポートフォリオ組成、運用に入ったあとの銘柄入れ替えは、レベル 4 人工知能が自動的に実行してくれることになる。

このようにして銘柄選定タスクは終了した。次の柱である投資家その他への説明タスク（④と⑤）について考慮すると、こちらはかなり厄介である。まず、投資家や規制当局へのリスク説明。レベル 4 人工知能で構築されたプログラムは、最終ユーザー（この場合は個人投資家）にとってはもちろん、開発者（この場合はファンド・マネージャー）にとっても、ニューラル・ネットワークの中身をのぞき込んで解読しようとしても、それがどのように機能しているかはみえない。銘柄選定の理由（「なぜその銘柄を選んだのか」）は、何千ものニューロンに埋め込まれ、数百もの複雑に連結されたレイヤーに展開される。最初のレイヤーで何らかの入力、例えば上場企業の過去 20 年分の当期純利益の数値や株価といった入力を受け取り、ニューロン内で、重みを掛け合わせ、閾値を減算し、次のレイヤーに伝達される。この処理が繰り返され、最終の出力が生成される。さらに、ここにバックプロパゲーションのような調整処理が加わって、プロセスはさらに複雑となる。

情報科学者のバイロン・ワラスは、「マルチ・レイヤーの複雑なニューラル・ネットワークは、「隠れた」レイヤーを内包しており、入力と出力の関連性を調べたり、解釈したりすることは大変難しい」と WIRED 誌のインタビューで述べている⁶²。また、マサチューセッツ工科大学で深層学習のアプリケーションが専門のトミー・ジャコラ教授は、「ごく小さなニューラル・ネットワークがあるとすれば、もしかす

⁶¹ 一般社団法人投資信託協会『投資信託の主要統計』2018年6月

⁶² Matt Burgess "Holding AI to account: will algorithms ever be free from bias if they're created by humans?", WIRED UK, 11 Jan 2016 <http://www.wired.co.uk/article/creating-transparent-ai-algorithms-machine-learning>

ると(ニューラル・ネットワークがその結果を出した理由を)解読することも可能かもしれない。しかし、レイヤーが数百にもなり、レイヤー当たりニューロンが数百にもなる大きなニューラル・ネットワークにもなると、人間には理解不能となる」と述べている⁶³。

この開発者がのぞきこんでもよくわからないニューロネットワークに内在するリスクは、リターンとトレードオフの関係にある従来のファイナンシャル理論で説明できるリスクではない。理論的にも実証的にもトレードオフの関係は、今のところ証明できておらず、市場メカニズムとは関係のないリスクである。いわば、単なるコンピュータシステムの「瑕疵」であり、流動性リスクなどの市場メカニズムに関わるリスクとは、リスクの所在も性質も全く異なる。投資家には、「ディープラーニング・ブラックボックス・リスク」とでも名付けた新しいリスク概念で説明するほかない。このリスクは、また、マーケット・リスクのように人間が考えたファイナンシャル理論をプログラムに落とし込んだものではないので、因果関係にもとづいた論理的定量化の方法をもたない。定量化できないリスクとは、予測不能とほぼ同義となる。

さて、この状態において、ファンドマネージャーが投資家から「なぜパフォーマンスの悪い銘柄ばかりを選んだのか」と問われたとき、ファンドマネージャーが投資家に返し得る、科学的に正しく、投資家にもわかりやすい回答は、「わからない」である。もちろん、モデルとデータの設計には関与しているので、「これこれこういう学習データを使って、こういうアルゴリズムでモデルを作り、バックテストではこれくらいの高いパフォーマンスを上げるところまで調整して、検証したのですが、実運用ではそのとおりには機能しませんでした」というような説明も、もちろん正しい。しかし、この説明でも、なぜ開発時の検証で達成できたパフォーマンスが、実運用で達成できなかったのかは「わからない」ことには変わらない。あるいは、苦し紛れに、「テーマ型ファンドなので市場の下落より振れ幅が大きくなりました」というような、論理設計に基づかない後付け説明でその場を乗り切ろうとすると、「先月マーケットが高騰したときは、マーケットほどには上昇しなかったではないか」と投資家から切り返されれば、答えに窮することになる。こんなことでは、世界中ほとんどの国の金融法制で義務付けられた、投資家保護のためのリスク説明責任を果たしているとは言い難い。

以上、簡単にファンドマネージャー機能をレベル4人工知能に置換した場合の仮定的・簡易的システム企画分析をおこなった。結論的に、レベル4人工知能によるファンドマネージャー機能の置換は、技術的には可能であるが、経済合理性、法令順守、ブランド信用力維持などの観点から、「やらない方がよい」と推奨するのが誠実なシステム企画者というものであろう。全面置換を難しくしていたのは、レベル4人工知能が原理的にもブラックボックス性、透明性の問題である。「なぜその結果を出したのか」を論理的に説明することができないレベル4人工知能は、重い説明責任が法的・社会的に課せられるファンドマネージャーの業務には向かないのである。このことが、レベル4人工知能に対して、ファンドマネージャーを注意深くさせているのであろう。

⁶³ Will Knight "The Dark Secret at the Heart of AI", April 11, 2017, MIT Technology Review May/June 2017 Issue

ハーバード大学と共同で、「責任ある人工知能の利用委員会 (the Council on the Responsible Use of Artificial Intelligence)」を立ち上げたバンク・オブ・アメリカの最高執行責任者(COO)で最高技術責任者(CTO)のキャシー・ベサントは、「人工知能の開発・販売業者が(人工知能の)モデルやデータソースの議論を独占している。これは、我々ユーザーからしてみると、中身がみえるかどうかわからないブラックボックスを作っているに他ならない。」とフォーブス誌のインタビューで述べ⁶⁴、レベル4人工知能が、金融機関自身にとってもブラックボックスとなってしまうことに警告を発している。レベル4人工知能のブラックボックス性、透明性の問題は、レベル4の実装が最も進行しているように見える金融業界の当事者にとってすら、大きな課題として社会に問題提起されているのである。

4.2.4 「疑似ソフト化されたハード」：トレーディング・プログラム

それでは、レベル4人工知能の実装が金融市場の4機能の中で最も進んだトレーダー機能においては、このレベル4のブラックボックス性、透明性欠如の問題にどう対処しているのだろうか。トレーダーあるいはトレーディング機能のシステム開発者たちは、レベル4のブラックボックス問題を克服できているのであろうか。

この問いに対しては、第一に、ファンドマネージャーとトレーダーの法制度的な責任範囲の違いにひとつの関連性を見出すことができる。ファンドマネージャーの最終ユーザーは個人である。投資信託の場合は、直接的に個人がユーザーとなり、年金の場合は年金基金を通じて間接的に、保険の場合は保険会社・保険商品を通じて、二重に間接的に個人が最終ユーザーとなる。ここには、個人に広く投資の門戸を開いているという公共性があり、ファンドマネージャーは個人投資家の財産を保護するという公共的使命・責任を負っているのである。一方、トレーダーの最終ユーザーは、ほとんどの場合、法人である。大きく2種類に大別でき、ひとつは一般事業法人や他の金融機関、ファンドマネージャーなどからの注文を取り次ぐトレーダーであり、もうひとつは、その金融機関自体の自己資産を使って取引するトレーダーである。

- 前者の取り次ぎの場合、顧客は法人であり、また、トレーダー自身は売買する銘柄や金融商品の選定には、多くの場合は関与せず、発注を取り次ぐ場合がほとんどである。「なぜその金融商品の売買を選択したのか」の説明責任は、顧客側の組織の問題であり、トレーダーは注文に従って最も条件のよい価格での取引(最良執行条件)の責任及びそのテクニックを問われるのみであり、説明責任はファンドマネージャーに比べてずっと軽い。
- 後者の自己資産を任される場合、自己資産なので、どのような手法で自己の資産の取引をしようとする(株価操作やインサイダー取引などの不正取引はもちろん除外して)、ブラックボックスであろうと、占星術であろうと、トレーダー個人の説明不能な直観であろうと構わない。法制度的にも、自己資産の投資は私的な行為とみなされるので、そこでの説明責任はトレーダーとその企業の経営者・株主との関係性において閉鎖的に決定して構わないということになる。他人の資産を預かるファンドマネージャーに比べて公共性は低く、制度的に課せられている説明責任は(不正取引の場合は除いて)ずっと軽い。

⁶⁴ “Time to break open AI's black box, and keep it open”, ZDNet, May 12, 2018

金融市場の中で、公共性が高く、重い説明責任が課されたファンドマネージャー機能では、レベル4人工知能の実装が限定的にしか進行しておらず、私的で説明責任の軽いトレーダー機能では、レベル4人工知能実装が最も進んでいる。このことは、レベル4人工知能に係わるひとつの問題、バンク・オブ・アメリカのCTO、キャサリン・ベサントが懸念を示したブラックボックス性の問題との関連性を示唆している。

トレーダー機能における透明性問題への対応について、第二に、トレーダー機能のレベル4人工知能の適用対象と適用方法から、もうひとつの関連性が見出せる。今回対象としたトレーダー機能のレベル4人工知能実装事例の適用対象とその適用対象の捉え方を掘り下げてみると、事例のすべてが、本来金融市場を捉えるべき人工抽象システム（ソフト2類）ではなく、機能対象を疑似的に自然・人工物理システム（ハード1類・ハード2類）に見立てるようにして、金融市場を捉えていた。その機能対象の捉え方には以下の4つのアプローチがあった。

- 金融市場における株価変動や為替変動を波形記憶として画像的に捉える方法。これは金融市場を疑似的に人工物理システム（ハード2類）としてみたてた捉え方である。「AI、投信・外貨預金で活躍―銘柄選択、レア情報もカバー(M&D)」(事例 NO:311)などが該当する
- 不正な株式・為替取引の発見に対して、発注状況(板)を波形記憶として画像的に捉える方法。これは発注状況を疑似的に人工物理システム（ハード2類）としてみたてた捉え方である。「AIで不正取引を、SBIや楽天証券、NECと実験」(事例 NO:328)などが該当する
- 金融市場のプレイヤーの生理的反応を捉える方法。これは金融市場の自然物理システム(ハード1類)として捉えるものである。例えば中央銀行総裁の声、表情などから言語には表れない政策発表のインサイトを読み取ろうとする事例、「金融政策変更の予兆、日銀総裁の表情に？、記者会見の映像、民間がAIで分析。」(事例 NO:166)が該当する
- 経済変動をアナリスト・レポートなどの資料の単語から特徴量を抽出する方法。これは経済を疑似的に人工物理システム（ハード1類）とみたてた捉え方である。「AI地域別に景気分析 大和総研がツール開発」(事例 NO:378)などが該当する

トレーダー機能における透明性問題への対応について、第三に、そもそも、ゴールドマン・サックスのトレーダーを激減させた主要因がレベル4なのかという疑問がある。実は、トレーダーをコンピュータに置換する動きは、1980年代から「プログラム・トレーディング」や「マシン・トレーディング」の名称でゆっくり進行していた。1987年のブラックマンデーは、プログラム取引に起因する最初の暴落であると考えられている。マシン・トレーディングは1990年代後半には「アルゴリズムック・トレーディング」の名称が一般的となり、2000年代になるとそこに超高頻度取引(High Frequency Trading)の要素が加わり、自動トレーディング・プログラムによる取引の比率がさらに高まった。しかし、自動トレーディング・プログラムによる取引の比率は、1980年代から今日までを通じて、概ね徐々に増加してきたもので、ディープラーニング・ブーム以降に飛躍したのではない。米国株式市場における超高頻度取引の比

率は2005年に21.8%、2009年に61%に達するが⁶⁵、現在も6割⁶⁶と2009年と同水準に留まっている。日本の株式市場では、2005年にはほぼゼロと推測できるが、2013年には注文金額の51%、売買代金の25.9%⁶⁷、現在では売買代金の3~4割⁶⁸と報告されており、日米いずれの市場においても第三次人工知能ブームが起こる以前に急激な伸びがあり、ブーム発生後は安定的、緩慢な伸びを示している。

質的な変化についても、同様のことがいえる。1980年代当初のマシン・トレーディングは、複数の機関投資家からの注文をひとつにまとめる程度の機能しかもたない補助的ツールであった。その後、1990年代後半には、離散時間市場モデルや Grossman-Miller モデル、板情報力学を説明変数に加えた確率論的モデル⁶⁹に進展していく。2000年代には、定量的なデータのみならず、ニュースなど定性的データの多変量解析モデルがアルゴリズム・トレーディングのエンジンに埋め込まれるなどの進展がみられた。しかし、近年の第三次人工知能ブーム以降、急激にレベル4人工知能が実装された傾向は特になく、アルゴリズム・トレーディング・エンジンが質的に変化した様子もみられない。今日の実運用されているアルゴリズム・トレーディング・プログラムの大部分も、レベル2/3人工知能によって構成され、レベル4は概ね製品化のための開発段階（ステージ4）にあり、一部補完的なパーツとして実装されているにすぎない⁷⁰。

このように自動トレーディング・プログラムは、量的にも質的にも、レベル4人工知能の急拡大とは、それほど関係なく、ゴールドマン・サックスのトレーダーの激減を主導したのはレベル4人工知能ではないといってよい。この項の初めに人間のトレーダーに求められる能力として、「論理的思考による判断・行動が、ほとんど反射的にできるような瞬発力」と述べた。それはまさにコンピュータに適した仕事であり、それゆえに1980年代からすでにコンピュータによる人間のトレーダーの置換が始まっていたのであるが、その30年の歴史の大部分を主導してきたのはレベル2/3人工知能であり、レベル4ではない。

4.2.5 民主主義の人工知能、権威主義の人工知能

ゴールドマン・サックスの前CIO/現CFO代行のチャベスがいうように、急速に人間のトレーダーがコンピュータによって置換される金融市場であるが、そこでのレベル4人工知能の関わりの詳細を調べた事例分析からは、次のことが明らかになった。

- 法制度的に説明責任負荷が最も軽いトレーダー機能において、人工知能による人間のトレーダーの置換が最も進行し、説明責任負荷がもっとも重いファンドマネージャー機能において、人工知能による人間のファンドマネージャー置換が最も遅かった

⁶⁵ World Federation of Exchange "World Federation of Exchanges Reports on High Frequency Trading (HFT)" Wednesday, 29 May 2013

⁶⁶ 松崎 泰弘 『ここまできた「超高速」株取引の威力：「HFT」は善か悪か？』 東洋経済新報社クオンツ・リサーチ 会社四季報 ONLINE 2017年07月20日

⁶⁷ 保坂豪 『東京証券取引所における High-Frequency Trading の分析』 JPX ワーキング・ペーパー 2014年5月

⁶⁸ 松崎 泰弘 『ここまできた「超高速」株取引の威力：「HFT」は善か悪か？』 東洋経済新報社クオンツ・リサーチ 会社四季報 ONLINE 2017年07月20日

⁶⁹ 足立高德 『アルゴリズム取引』 朝倉書店 2018年06月

⁷⁰ Andrés Arévalo, Jaime Niño, German Hernández, Javier Sandoval, "High-Frequency Trading Strategy Based on Deep Neural Networks", Conference paper: International Conference on Intelligent Computing, ICIC 2016: Intelligent Computing Methodologies pp 424-436, 12 July 2016

- 人工知能による人間機能の置換が最も進行しているトレーダー機能においても、本来人工抽象システム（ソフト2類）である金融市場を疑似的に自然・人工物理システム（ハード1/2類）として捉え、人工抽象システム（ソフト2類）への真正面からの適用を「回避」しようとする実装傾向が顕著にみられた
- 人間のトレーダー機能の人工知能への置換は、30年の間を通じて進行しており、この数年間にレベル4人工知能によって急激に起こったものではなく、現在の自動トレーディング・プログラム内でも、レベル4は補完的にしか使われていない

これらの事象から示唆されるのは、レベル4人工知能が、金融市場という人工抽象システム（ソフト2類）において、あたかも、その肝心な中核部を「回避する」かのように進行してきているということであり、そして、レベル4人工知能が「回避」してきた肝心の中核部にあるのは「説明責任」ではないかということである。

伝統的なコンピューティングが人間の考え出した規則や論理をプログラムに書き込んで機能するのに対して、レベル4人工知能は（学習データとアルゴリズムの設計に人間が関与するもの）そのデータから言語化されていない「特徴」を見出す。出力された結果について、なぜその結果を出したのかは、人間にはその開発者にとってすらほとんどわからない。人工抽象システム（ソフト2類）で成立している金融市場にとって、それを民主的に投資家利益保護の理念から公明正大に運営しようとするとき、「説明責任」は本質的に重要な成立要件であり、そこでは、レベル4人工知能の「透明性」欠如の問題は、致命的な欠陥となり得る。前述のMITのジャコラは、「（ディープラーニングの）透明性（transparency）」の問題は、「今日すでに重要な問題であるが、将来さらに重要な問題となる」と述べている⁷¹。

ここでは金融市場を事例として取り上げたが、説明責任・透明性は金融市場に限らず、広く人工抽象システム全般に通じる基盤的な要素である。とりわけ、民主主義社会にとっては、説明責任はその成立要件であるといってもよい。政治学者のフランシス・フクヤマは、近代的な自由・民主主義(liberal-democracy)の基礎は3つの制度によってできあがったと述べている。第一に「国家」（国家は権力を統合し、区切られた領域内で執行することができる）、第二に「法の支配」（権力は法のもとで正当化される一方で制限も受け、権力行使が予見可能なものとなる）、第三に、民主主義的な「説明責任」（accountabilities）（国家権力が支配者の利害だけでなく共同体全体の利害を反映することを確かなものにできる）⁷²とし、「説明責任」を自由主義・民主主義存立要件のひとつであると位置づけている。

現実の民主主義社会に目を向けると、レベル4人工知能は、民主主義社会の市民の投票行動にも大きな影響を及ぼすようになってきている。例えば、Twitter上ではbotとよばれる人工知能による自動Twitterが、全体のツイート・トラフィックの約半分を占めるようになってきているとの報告もあり⁷³、そのうち英米の政

⁷¹ Will Knight "The Dark Secret at the Heart of AI", April 11, 2017, MIT Technology Review May/June 2017 Issue

⁷² フランシス・フクヤマ 『政治の起源：人類以前からフランス革命まで』 講談社 2013.11

⁷³ クラウド運営サービス会社のインカプラ社の調査による：Incapsula "Bot Traffic Report 2016"、「我々のネットワークにおいて」と調査対象範囲を明示しないデータが示されている。 <https://www.incapsula.com/blog/bot-traffic-report-2016.html>

治 bot に関しては、約 4 割が bot で残り約 6 割が人間のツイートであるという報告もある⁷⁴。さらに、英国スウォンジー大学のオレクサンダー・タラベラらは、この Bot による政治ツイートが、2016 年のイギリスの Brexit 国民投票では離脱派票を 1.76% 押し上げ、同年のアメリカ大統領選挙ではトランプ票を 3.23% 押し上げたとの研究結果を報告している⁷⁵（引用は同研究論文を報じたブルンバーグの記事⁷⁶から）。

また、2017 年のイギリス総選挙、2016 年のアメリカ大統領選挙やその他の国の主要な選挙において、政治 Bot とソーシャル・ネットワークの「コンテンツ・パーソナライゼーション・システム」（Facebook のニュースフィードなど）は、共振的に、フェイクニュースを拡散させた。Facebook は大統領選後すぐ 2016 年 11 月下旬には、フェイクニュース拡散防止策として、フェイクニュースを検知してユーザーに推奨しない「よりよいシステム」を導入すると発表した⁷⁷が、2017 年夏には、人工知能によるフェイクニュース検知を諦め、10,000 人の人間のアナリストによる人海戦術に置き換えている⁷⁸。また、ソーシャル・ネットワークのコンテンツ・パーソナライゼーション・システムは、同じ政治的傾向の人々を「集める」ように作用するため、「意見の多様性を阻害し、政治的な透明性（disclosure）を弱体化させてしまう」（オックスフォード大学インターネット研究所(OII)のブレント・ミッタールスタッド）⁷⁹との指摘もされている。

日本経済新聞は、第 2 章で事例として取り上げた記事（事例番号 738）の中で、「文化大革命時代の中国で地主の家に生まれ育ったためにつり上げの標的にされた経験をもつ SF 作家の王晋康氏が、人工知能（AI）と出会い、「AI の決定の方が（過ちに気付かない人より）信じられるかもしれない」と感じている」とのエピソードを引き合いにだしながら、シルバー民主主義の政治危機に襲われる日本でも、「私心」を捨てられる）AI を政策決定に取り入れるべきではとの問題提起をおこなっている⁸⁰。

フクヤマは、中国を人類史上最初に「国家」（自由民主主義の第一の制度的要件）を成立させた社会であると位置づけているが、現代の中国は残りの 2 要件「法の支配」と「説明責任」を充足しないと位置づけている。英エコノミスト誌のシンクタンク部門、EIU の民主主義指標においても、中国は、最も民主主義度が低い「権威主義」と評価されている（EIU 民主主義指標のカテゴリは以下のとおり「完全な民主主義」「不完全な民主主義」「民主-権威主義ハイブリッド」「権威主義」）⁸¹。日本は、2014 年まで「完全な民主主義」と評価されていたが、2015 年に「不完全な民主主義」に転落したものの、「不完全な民主主義」カテゴリ内の上位層には踏みとどまっている⁸²。

⁷⁴ Yuriy Gorodnichenko, Tho Pham, Oleksandr Talavera, "Social media, sentiment and public opinions: Evidence from #Brexit and the #USElection", VoxEU.org, The Centre for Economic Policy Research Policy Portal, 02 June 2018

⁷⁵ Yuriy Gorodnichenko, Tho Pham, Oleksandr Talavera, "Social media, sentiment and public opinions: Evidence from #Brexit and the #USElection", VoxEU.org, The Centre for Economic Policy Research Policy Portal, 02 June 2018

⁷⁶ "Twitter Bots Helped Trump and Brexit Win, Economic Study Says" Jeanna Smialek, BLOOMBERG, 2018年5月 21 日

⁷⁷ "Zuckerberg reveals plans to address misinformation on Facebook", Kate Conger, TechCrunch, Nov 20, 2016

⁷⁸ Polonski 2017

⁷⁹ Brent Mittelstadt "Auditing for Transparency in Content Personalization Systems", International Journal of Communication 10 (2016)

⁸⁰ 日本経済新聞 「気がつけばそこに（4）政治の限界を超えて 私心捨てられるか」 2017/2/2（事例番号 738）

⁸¹ The Economist Intelligence Unit "Democracy Index 2017"

⁸² The Economist Intelligence Unit "Democracy Index 2014", "Democracy Index 2015", "Democracy Index 2016", "Democracy Index 2017"

その中国では、レベル4人工知能の社会実装が世界をリードするほどに進行している。例えば顔認証の技術の分野では、顔認証技術を駆使した「全能社会信用システム」のパイロット・プロジェクトを2015年に持ち上げ、2020年までには、些細な交通ルール違反なども含め、全国民の「信用度」を格付けするシステムを実装するという⁸³。2018年4月には、顔認証システムにより、コンサート会場60,000人の中から経済犯を見つけ出して逮捕している⁸⁴。2019年開港予定で建設が進む北京大興新国際空港には、顔認証により個人を特定して、空港にいるすべての人間の行動を監視するシステムが導入されるほか、このシステムの請負業者の1社である人工知能スタートアップ、依图技術(Yitu Technology)によると、内容を一切話せない機能をもったシステムも導入されるという⁸⁵。また、中国政府はすでに「罪を犯しそうな」人間を「罪を犯す前に」検知するシステムを導入している⁸⁶。顔認識技術は、すでに人間個々人の性的指向、政治的指向、知能指数までも顔の特徴から検知できるようになっており⁸⁷、そうした技術が中国などで導入されていくのは、時間の問題かもしれない。

政治 Bot やフェイクニュースの拡散が、重要な選挙結果にも影響を及ぼすようになった日本や欧米などの民主主義社会であるが、そこには民主主義を維持するための健全な批判が活発に議論され、人工知能の倫理を構築しようとする社会の努力がある⁸⁸。一方で、国家権力が、国民監視のためのレベル4人工知能の実装を急速に進める権威主義国家の中国がある。レベル4人工知能を人工抽象システム(ソフト2類)に適用しようとするときの「透明性」の問題は、民主主義と権威主義という、この2つの異なる社会では異なる意味をもつ。前者にとっては、社会が解決すべき挑戦的課題、後者にとっては権威主義国家権力の維持にとって都合ということであろう。バンク・オブ・アメリカのCTO キャシー・ベサントが懸念を示したように、人工知能の開発業者・販売業者たちが人工知能の議論を独占し、レベル4人工知能の人工抽象システム(ソフト2類)への適用を、透明性・論理性・説明責任をないがしたまま社会実装を進めるようなことがあっては、透明性と論理性によって支えられた自由主義・民主主義社会の根幹を揺るがすことになりかねない。

4.3 創発性の問題

4.3.1 ピークは過ぎた FinTech、次のブームを狙う? 「HR テック」

日立製作所は、2016年12月、人事の分野に人工知能を適用したサービス、「リシテア/AI分析」を2017年2月に販売開始すると発表した。ビッグデータや人工知能を人事の分野に応用して、個人と組織のパ

⁸³ "China 'social credit': Beijing sets up huge system", Celia Hatton, BBC News, Beijing, 26 October 2015

⁸⁴ "Chinese man caught by facial recognition at pop concert", BBC News, 13 April 2018

⁸⁵ "Beijing's new Zaha Hadid-designed airport to showcase latest facial recognition technology", South China Morning Post, 18 July 2018

⁸⁶ "I was shocked it was so easy": meet the professor who says facial recognition can tell if you're gay" Paul Lewis, The Guardian, 7 Jul 2018

⁸⁷ "I was shocked it was so easy": meet the professor who says facial recognition can tell if you're gay" Paul Lewis, The Guardian, 7 Jul 2018

⁸⁸ 新保史生「ロボット・AIと法をめぐる国内の政策動向」 特集「AI社会論」 『人工知能』 VOL 32 NO.5 (2017年9月号) 人工知能学会

パフォーマンス最大化に向けた診断や予測を提供し、新たな価値を創出するサービスであるという⁸⁹。2017年9月からは、この「リシテア/AI分析」を使い、（リクルートやデロイトなど日系・外資系大手コンサルティング・ファームが席卷する）人事コンサルティングの分野に参入すると日本経済新聞の取材に対して述べている。

日本経済新聞は「リシテア/AI分析」の機能について、「HRテック」（human resource technology）との一般呼称に分類しつつ、次のように報道している。「社員に関する適性試験や面談などの結果を数値化。AIで分析し、「責任感が強い」「大局観を持つ」「勝負志向」といった特性ごとに社員を分類する。社内にどのようなタイプの人材がいるか分布がわかるようになる。安定重視の社員が多い企業には、勝負志向が強い人材の採用を増やすよう勧めて採用計画に役立ててもらおう。社員の人物像や特性を具体的なデータとして可視化できる。経験や感覚に頼っていた従来の人事手法に比べ組織の実態を把握しやすくなり、採用や育成の計画立案が容易になる。日立は社内での活用例も増やし、蓄えたデータを解析に生かす。」⁹⁰

「リシテア/AI分析」にレベル4人工知能が使われているかについては、公表された情報からは明確な判定はできない。しかし、広く世界に目を向けると、レベル4人工知能を人事的意思決定の分野に適用した「HRテック」はかなり製品化が進んでいる。

- 東京・渋谷に拠点をおく人材サービスのビズリーチは、「深層学習（ディープラーニング）により、AI（人工知能）が蓄積した情報を分析、最適な人材を選ぶ基準をはじき出す評価ツール」を2017年春から開始し⁹¹、同年秋からは、これに勤怠管理機能を加え、評価ツールと合わせて顧客企業への導入を開始している⁹²。「採用面接から現在に至る評価の積み重ねだけでなく、働きぶりを追跡調査してデータベースにする。これを、大量のデータから特徴を自分で見つけて学習するAIの基幹技術、深層学習（ディープラーニング）で分析。評価や最適な職場、ポジションをはじき出す」⁹³という。

- 米国・サンフランシスコ拠点のベンチャーMyaSystemは、「ディープラーニング・ベースの複数意図分類器（multiple intent classifier）、固有表現抽出（named entity extraction：NEE）及び文章意味解析（sentence semantic analysis）により構成される」⁹⁴採用面接チャットボット（recruitment chatbot）サービスを提供している。応募者の面接をしたあと、職務経歴書

⁸⁹ 株式会社日立ソリューションズ・プレスリリース：

<https://www.hitachi-solutions.co.jp/company/press/news/2016/1201.html>

⁹⁰ 事例番号244：日本経済新聞2017/09/22記事「日立、人事コンサル参入、AIで適正判断、採用助言」

⁹¹ 日本経済新聞2016/6/14「ビズリーチ、人事管理にAI導入 採用から配置まで一元管理」

https://www.nikkei.com/article/DGXLASDZ14H5T_U6A610C1000000/

⁹² 日本経済新聞2016/6/5「AIで人事部いらず？ データで最適配置 — ビズリーチやヤフー、人との役割分担探る」

<https://www.nikkei.com/article/DGXLZO03611460U6A610C1TI1000/>

⁹³ 日本経済新聞2016/6/5「AIで人事部いらず？ データで最適配置 — ビズリーチやヤフー、人との役割分担探る」

<https://www.nikkei.com/article/DGXLZO03611460U6A610C1TI1000/>

⁹⁴ <https://hiremya.com/#features>（MyaSystem Web サイト）

やLinkedInのプロフィールと応募後の変更その他 Web 情報を総合して解析され、ランク付けされた採用候補者リストを顧客企業の（人間の）採用担当者に提出する⁹⁵。

- IBM、Apple、Amazon、ボーダフォン、インテルなど、グローバル企業 700 社以上で導入され、日本では三井不動産リアルティでの採用が決定⁹⁶している HireVue は、人工知能が動画面接するサービスで、人工知能からの質問に対する表情や言葉のパターンを解読し⁹⁷、採用者に合否判定結果を提供する。

- アイルランド拠点のスタートアップ、Opening.io は、自然言語処理、深層学習（教師付き及び教師なし）、検索エンジン Lucene の組み合わせによる職務経歴書(CV)と職務説明書(job description)のマッチング・サービスを API により提供している⁹⁸。

このように HR テックの社会実装が急速に進む中、その意義について、ビズリーチの南壮一郎社長は、「これまで人事評価は（上司の）勘や経験に頼ることが多かった」と指摘し、日本経済新聞は、「デジタル革命が多く部門で進むなか、勘や経験をベースにしたアナログな判断も目立つ人事部。好き嫌いで評価が左右され、社員が士気を下げてしまうことだって少なくない。だからこそ AI 人事の需要は大きい」⁹⁹と（筆者匿名の記事で）論評している。また、慶應義塾大学ビジネス・スクールの岩本隆特任教授は、深層学習の発達により HR テックのスタートアップへの資金投入や案件件数が 2-3 年に急激に伸びており、これにより、「経営や人材マネジメントにおいて HR テクノロジーが寄与する部分が多くなるため、HR テクノロジーをどう活用するかは企業の競争力を高めるうえでますます重要になる」¹⁰⁰と述べている。

本節では、レベル 4 人工知能の実装が進む HR テックを事例として取り上げ、その深掘りの分析からレベル 4 人工知能のソフト 3 類への適用の問題を考えていく。

4.3.2 「アルゴリズムック・バイアス」－偏見を内蔵する人工知能

レベル 4 人工知能の実装が進む HR テックであるが、そこには、意図せずした雇用差別をもたらす危険性が潜むことが、いくつかの実証研究により示されている。例えば、ロンドン・ビジネス・スクールのアンジャ・ランブリヒトと MIT スローン経営大学院のキャサリン・タッカーは、STEM (science, technology, engineering and mathematics) 系職種の求人広告を Facebook、Google AdWorks (Google の広告ネットワーク)、Instagram、Twitter を通じて世界 191 カ国に配信するテストをおこなっている。二人の実験結果報告によると、4 つの広告ネットワークすべてで、女性より男性に多く広告が配信されるという「差別的な結果」(discriminatory outcomes) となったという。その差は、最も男女差の小さい Twitter でも男性への配信量 (impressions) が女性への配信量より 26.5%多く、ほぼ、同レベルの Facebook が男性

⁹⁵ "Chatbots increase recruitment opportunities", Business Insider Intelligence Jul. 12, 2016, 10:30 PM

⁹⁶ タレント株式会社広報 「三井不動産リアルティ株式会社によるデジタル面接プラットフォーム「HireVue」導入決定のお知らせ」 2018年6月6日 10時00分

⁹⁷ The Economist, Jun 28, 2018, "How an Algorithm May Decide Your Career Getting a job means getting past the computer" <https://www.economist.com/business/2018/06/23/how-an-algorithm-may-decide-your-career>

⁹⁸ 同社 Web サイト: <http://www.opening.io/match>

⁹⁹ 日本経済新聞 2016/6/5 「AIで人事部いらす？ データで最適配置 - ビズリーチやヤフー、人との役割分担探る」 <https://www.nikkei.com/article/DGXLZO03611460U6A610C1TI1000/>

¹⁰⁰ 日経 BigData 2017/7/7 「将来の活躍や離職、AIで予測 「HR テック」で人材管理」 <https://www.nikkei.com/article/DGXMZO17003950Z20C17A5000000/>

への配信量が女性への配信量を 27.6%上回り、最も大きい Instagram では男性への配信量は女性への配信量の 5.15 倍にもものぼっていた¹⁰¹。

二人は、この実証研究の結果を整理した論文の中で、この「雇用差別的な結果」が何に起因するのかを詳細に分析し、想定しうるいくつかの要因をひとつずつ丁寧に検証している¹⁰²：

(a) (雇用差別的な結果は) SNS/Web ユーザー人口の男女差によるものか? : そうではない。例えば、Facebook の全世界のアクティブ・ユーザー人口の男女比は、41:59 で女性の方が多い。

(b) (雇用差別的な結果は) 女性の STEM 職に対する関心が男性より低く、それに関わる人間行動を人工知能が学習したことによるものか? : そうではない。4 つの広告ネットワークすべてにおいて、STEM 職求人広告のクリック率は女性の方が男性より高かった (女性の方が男性より STEM 職への関心が高いことを示している)。

(c) (雇用差別的な結果は) 男女の雇用差別意識の強い社会があり、それに関わる人間行動を人工知能が学習したことにより起因するものか? : そうとは言えない。二人は、世界銀行のジェンダー・データ・ポータルに掲載されているデータから労働市場における男女格差を表す指標を作成し、国別広告配信データと突き合わせてみたが、二人の作成した男女雇用格差指標と STEM 職求人広告の配信男女差との関連性はみられなかった。

(d) (雇用差別的な結果は) 広告ネットワークの配信アルゴリズム内に、求人広告に関わる (意図的な) 男女差別が組み込まれているからか? : 少なくとも Facebook については、そうではないと推測される。二人は実験的に Facebook に性別を特定したダミーの求人広告を送付したところ、「差別的なので配信できない」旨の拒否通知があり、配信されなかった。(他の広告ネットワークでは同様のテストは行っていない。)

二人はこのように想定され得る諸要因を排除したうえで、雇用差別的な結果を引き起こす原因として、広告ターゲットとしての女性層の ROI (投資対費用効果) の低さ (コストの高さ) に見い出した。女性層、とりわけ 25-44 歳の女性層は、家具やファッションなどの高級消費財の主要購買層であり、それらの商品の広告主が高値を付ける。STEM 職の求人広告は、それらの小売業者との個々の広告選定自動オークションにおいて、「負けている」場合が多いことが、差別的な結果を引き起こした原因と結論付けている¹⁰³。

しかし、この結論には、人工知能研究会の視点からは、一点の問題提起をしておかなくてはならない。STEM 職求人がファッション広告に負ける率が女性の方が高いことこそが、需要 (求人) と供給 (応募候補者の関心) のバランスにおいて、女性の STEM 職への関心が、ファッションや家具などへの関心に比べて男性より相対的に低いことの経済合理的な結果ととらえ得る。したがって、女性の勝率が低いことは、

¹⁰¹ Anja Lambrecht, Catherine E. Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads", Social Science Research Network (SSRN) Electronic Journal, January 2016

¹⁰² Anja Lambrecht, Catherine E. Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads", Social Science Research Network (SSRN) Electronic Journal, January 2016

¹⁰³ Anja Lambrecht, Catherine E. Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads", Social Science Research Network (SSRN) Electronic Journal, January 2016

二人が上げた上記の反証(a)(b)を自身で反証してしまっていることになる。そのうえ、特に Instagram の 5 倍を超える男女差は、この推論のために二人が提示した男女の広告価格のわずかな差では、十分に説明できていない。(この点については、Facebook の広告ネットワークと Instagram の広告ネットワークにおけるレベル 4 人工知能の構成比率が興味深い示唆を与えている。Facebook の広告ネットワークは、Facebook 上のユーザーの行動情報(クリック、いいね、コメント、シェア、友達、アウトバウンド・クリック、フォロー、ブロック/非表示)を入力し、適合スコア(relevance score)を出力する決定木やロジスティック回帰などレベル 3 人工知能主体、補完的にニューロネットワーク(レベル 4)を統合したアルゴリズムで構成されている¹⁰⁴。それに対して、Instagram は、Facebook による買収後は Facebook の広告ネットワークを利用しているものの、元々、ビジュアル系に差別化されたソーシャル・ネットワークであるため、広告ターゲティングのもととなる特徴量抽出のアルゴリズムは、レベル 4 人工知能を主体として構成されている¹⁰⁵。このことが、Facebook の比較的小さい配信量の男女差(26.7%)と Instagram の極端な配信量の男女差(5.15 倍)と関連しているのかもしれない。)

いずれにしても、二人の実験データにより、人事採用意思決定プロセスの一部に人工知能のアルゴリズムを通すことによって、意図しない形で男女差別的な結果をもたらすことは示された。論文は、「明白な『アルゴリズムック・バイアス』(人工知能に内在する差別的なバイアスのこと)は起こり得る。しかも、それは、意図せずして、中立性を意図して設計された中立的アルゴリズムにおいてすら、起こり得る」¹⁰⁶と締めくくられている。

人工知能による意図しない差別の危険性を示した実証研究は、他にも多く報告されている。例えば、バージニア大学コンピューター・サイエンス学部のビンセント・オールドネスらのチームは、Facebook とマイクロソフトが AI 開発者向けに公開しているオープンデータに性差別的バイアスが内蔵されているかどうかを検査した。チームがバイアスを検査したのは、Facebook が意味役割付与(visual semantic role labeling)加工した画像データセット(imSitu vSRL)とマイクロソフトのマルチラベル分類(multilabel object classification: MLC)のデータセット(Microsoft COCO MLC)で、どちらのデータセットもレベル 4 人工知能をベースとした意味役割付与により生成されている。チームは、imSite データセットの 125,000 イメージのうち人間が登場する約 60,000 イメージを畳み込みニューラル・ネットワーク(Convolutional neural network: CNN)の特徴量によって抽出された状況、活動、役割、名詞を、独自に開発した人間の論理的思考を構造化したバイアス計測モデル(レベル 3 人工知能 CRF で構成されている)で解析することにより畳み込みニューラル・ネットワークの抽出した結果にジェンダー・バイアスが

¹⁰⁴ 2016 年 5 月時点の Facebook ニュースフィードに関するエンジニアリング・マネージャー、エワ・ドミノスカのプレゼンテーションによる。Ewa Dominowska, Engineering Manager, Facebook, "Generating a Billion Personal News Feeds", presentation at Machine Learning Conference, Seattle, May 2016, <https://youtu.be/iXKR3HE-m8c>

¹⁰⁵ Bernard Marr "The Amazing Ways Instagram Uses Big Data And Artificial Intelligence", FORBES, Mar 16, 2018

¹⁰⁶ Anja Lambrecht, Catherine E. Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads", Social Science Research Network (SSRN) Electronic Journal, January 2016

含まれているかどうかを検証した。また、マイクロソフトのMS-COCOについても、同様にバイアス計測モデルでジェンダー・バイアスを検証している¹⁰⁷。

その結果報告によると、Facebook imSitu vSRL と Microsoft COCO MLC のいずれのデータセットにおいても明白な性差別的バイアスがみられたという。例えば、Facebook の imSite vSRL データセットの場合、「料理 (cooking)」「掃除 (cleaning)」「洗濯 (washing)」「電子レンジ (microwaving)」「ショッピング (shopping)」「給仕 (serving)」などの行為には「女性」にバイアスがかかり、「指導 (coaching)」「調整 (tuning)」「目的達成 (aiming)」などの行為には「男性」にバイアスがかかっていた。このうち、例えば、「料理」という行為に対しては、「女性」が付与されている場合が、「男性」の2倍にもなっていた。また、MS-COCO MLC データセットの場合、「フォーク」「ナイフ」「スプーン」「ワイングラス」などの対象には「女性」にバイアスがかかり、「スノーボード」「バイク」「ボート」「スキー」「信号」などの対象には「男性」にバイアスがかかっていた¹⁰⁸。Facebook imSitu vSRL も Microsoft COCO MLC もオープンデータであり、世界中のレベル4人工知能の開発に利用されている。このことは、両データセットに内在する差別的なバイアスが世界中の人工知能アプリケーションに拡散している可能性を示唆している。

4.3.3 得意分野は時代遅れの恐怖モジュール

人工知能のアルゴリズムに内在するバイアスについては、他にも数多く報告されており、そのことを危惧するメディアやアカデミアからの論点提起も数多い。このことについて、英国バース大学のコンピュータ・サイエンティスト、ジョアンナ・ブリソンは、「多くの人がそれを人工知能の問題と指摘する。しかし、それは間違いだ。偏見をもっているのは私たち人間であり、人工知能は我々から偏見を学びとっているだけである。」¹⁰⁹ とガーディアン紙の科学記者、ハンナ・デブリンに語っている。

人間の意識の中には特定の人間の属性（人種、民族、性別、年齢など）に対する偏見やバイアスが潜んでいる。このことは多くの心理学実験によって確認されている。そのひとつ、おそらく最もよく用いられる潜在的な意識を測定する心理テスト、潜在連合テスト (implicit association test: IAT) による心理実験では、次のような潜在的な偏見・バイアスが確認されている¹¹⁰。

- 性差別に関連する偏見/バイアス：
 - 文系的職業 (artistic and humanities occupations) は「女性」に属し、理系的職業 (mathematics and engineering professions) は「男性」に属する
 - 「男性の名前」は「職業 (career)」に属し、「女性の名前」は「家庭」 (family) に属する

¹⁰⁷ Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, Kai-Wei Chang, “Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints”, Conference Paper, Conference: Empirical Methods in Natural Language Processing. EMNLP 2017. Copenhagen, Denmark. September 2017

¹⁰⁸ Zhao et al 2017

¹⁰⁹ Hannah Devlin “AI programs exhibit racial and gender biases, research reveals”, The Guardian, posted on Thu 13 Apr 2017 19.00 BST, Last modified on Wed 14 Feb 2018 21.11 GMT

¹¹⁰ Aylin Caliskan, Joanna J. Bryson, Arvind Narayanan, “Semantics derived automatically from language corpora contain human-like biases”, Science 356, 183–186 (2017) 14 April 2017, AAAS

- 年齢差別に関連する偏見/バイアス：
 - 「若さ」は「心地よい」（贈り物、幸せ、花、楽器などと同類）に属し、「年寄」は「不快」（昆虫、武器などと同類）に属する
- 人種・民族差別に関連する偏見/バイアス
 - ヨーロッパ系の氏名は「心地よい」に属し、アフリカ系の氏名は「不快」に属する（調査対象：アメリカ人、イギリス人）
 - 韓国人名は「心地よい」に属し、日本人名は「不快」に属する（調査対象：韓国系アメリカ人（ワシントン大学の学生））¹¹¹
 - 日本人名は「心地よい」に属し、韓国人名は「不快」に属する（調査対象：日本在住の日本人（名古屋大学の学生）¹¹²、日系アメリカ人（ワシントン大学の学生）¹¹³）

なお、このうち、人種・民族差別的バイアスについては、心理実験だけでなく、実際に企業が出した求人広告に応募するフィールド・ワークにおいてもバイアスが認められ、人種・民族差別的な偏見が実際の雇用差別となって表出していることも確かめられている¹¹⁴。

これら人間に対する心理実験やフィールド・ワークを踏まえ、「偏見をもっているのは私たち人間」と述べたブリソンは、プリンストン大学のコンピュータ・サイエンティスト、アービンド・ナラヤナンらとともに、人間の潜在的な差別的偏見やバイアスが人工知能のアウトプットにもみられることを確かめた。ブリソンのチームが調べたのは、Amazon が提供するオープンデータ、Common Crawl Corpus で、ここには過去7年間の世界中50億を超えるウェブページから収集されたウェブ・クロール¹¹⁵・データが格納されている。チームは、ここから引き出した8400億のトークン（大雑把に言って「単語」）を分析対象に、単語埋め込みベクトル・ツール GloVe (Global Vectors for Word Representation) をベースに、独自に開発した単語埋め込みツールを通じて、300次元への単語埋め込みをおこなった。さらに、Google News のビッグデータを分析対象に、同様な GloVe ベースの単語埋め込みツールを通じての単語埋め込み作業をおこなっている。

その結果、人間の潜在連合テストで検出されたレベルと同様の性差別的バイアス、年齢差別的バイアス、人種差別的バイアスが検知されたという。例えば、若さを心地よいものとし、年寄を不快なものとする年齢に関わる偏見／バイアスは、人間の潜在連合テストでは効果量 1.42（有意性 $(p) < 10^{-2}$ ）、ビッグデータの人工知能による分析では効果量 1.21（有意性 $(p) < 10^{-2}$ ）であった。また、ヨーロッパ系の氏名を心地よいものとし、アフリカ系の氏名を不快なものとする人種偏見／バイアスは、人間の潜在連合テスト

¹¹¹ Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz, "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test", *Journal of Personality and Social Psychology* Vol.74 No.6 1464-1480, 1998

¹¹² Miriam Steel, Sarah Teige-Mocigemba, "Cultural Adaptation (Im)Possible? Measuring Prejudice in Japan with the Affect Misattribution Procedure", *Psychologia*, 2014, 57, 201-212

¹¹³ Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz, "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test", *Journal of Personality and Social Psychology* Vol.74 No.6 1464-1480, 1998

¹¹⁴ Marianne Bertrand, Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", *American Economic Review* vol. 94, no. 4, September 2004

¹¹⁵ クロール：インターネット上のあらゆる Web サイトの情報を自動巡回プログラムが取得して、検索用データベース・インデックスを作成すること

では効果量 1.35 ($p < 10^{-5}$)、ビッグデータの人工知能による分析では効果量 1.41 ($p < 10^{-8}$) であった。ブリソンらは、「人間の潜在連合テストと人工知能のアウトプットが直接的に対応関係にあると主張するつもりはない」としながらも、両方の結果に「一様に高い (uniformly high)」バイアスが検知されたと結論づけている¹¹⁶。

文化人類学者クロード・レヴィ=ストロースが指摘するように、人類の多くの集団にとって、人間は自分の種族だけ、あるいは、同じ言葉話す言語集団だけ、ときには村の境界までとされる。そしてそれが高じて、自集団のメンバーを「人間」と呼び、他集団を野卑な「サルども」のような生き物とみなすのである¹¹⁷。何百万年もの間、人類は自分の社会集団を外敵から守る生存上の必要性に迫られていた。もちろん、有毒な動物や捕食者は恐ろしい外敵ではあったが、長い狩猟採集時代から数千年前までの人間集団にとって、第一次的な外敵は他の社会集団の人間であった¹¹⁸。先史時代の人骨から得られた考古学的証拠の集計結果によると、発掘された人骨のうち人間による殺傷の痕跡が確認できる人骨が、平均で約 15%、最も多い発掘地点では約 60%であった¹¹⁹。日本でも、縄文時代の殺人による死亡率は 1.81%、弥生時代初期には 50%、弥生時代全期通じて 3.62%と報告されている¹²⁰。このような状況下においては、他集団の人間（「よそ者」）に対する警戒心が、生存上の重要な資質となり、警戒心の強い個体がより寿命を延ばし、より多くの子孫を残していく（自然選択される）ことになる。

神経科学者のディーン・ブオノマーノによると、例えば、ガンのひなが、頭上を飛ぶタカの形状を識別してうずくまるといった防御行動を生得的に身につけていることや、チンパンジーが相当な「よそ者恐怖症」で、他の群れのオスに対する強い警戒心と攻撃性を発揮することなどにみられるように、「恐れ」は、動物が命を脅かす危険に対して確実に先回りして反応できるように適応、進化したものである。同様に、人間にも、（チンパンジーであったころから国家が出現する数千年前くらいまでの何百万年間）外敵から身を守るという適応課題への対応として、生存を脅かす最大の敵である「よそ者」を識別し恐れる性向が、ニューロンのネットワークに翻訳されて遺伝子に組み込まれてきたのだという¹²¹。

今日の私たちの「よそ者」を恐れる性向、その結果としての「よそ者」に対する偏見の抱きやすさには個人差があることがわかっているが、女性より男性の方が、若者より年長者の方が偏見を抱きやすく、実体的なものにも象徴的なものにも「恐れ」を抱きやすい人ほど、他集団の人間に偏見を抱きやすい。そのことをもって、狩猟採集時代の適応としての「恐れ」と今日の私たちの心の中の潜在的偏見の関連性を示す進化心理学的根拠であると主張する報告もある¹²²。また、偏見につながる「恐れ」の抱きやすさ（脅威

¹¹⁶ Aylin Caliskan, Joanna J. Bryson, Arvind Narayanan, “Semantics derived automatically from language corpora contain human-like biases”, *Science* 356, 183–186 (2017) 14 April 2017, AAAS

¹¹⁷ クロード・レヴィ=ストロース 『人種と歴史』 1981年 みすず書房

¹¹⁸ Richard B. Lee, Richard Daly “The Cambridge Encyclopedia of Hunters and Gatherers”, Cambridge University Press, 1999

¹¹⁹ スティーブン・ピンカー 『暴力の人類史』 青土社 2015/1/28

¹²⁰ Tomomi Nakagawa, Hisashi Nakao, Kohei Tamura, Yui Arimatsu, Naoko Matsumoto and Takehiko Matsugi “Violence and warfare in prehistoric Japan”, *Letters on Evolutionary Behavioral Science* Vol.8 No.1, 8-112017.4

¹²¹ ディーン・ブオノマーノ 『バグる脳：脳はけっこう頭が悪い』 河出書房新社 2012.12

¹²² Serdar Kaya “Outgroup Prejudice from an Evolutionary Perspective: Survey Evidence from Europe”, *Journal of International & Global Studies* VOL.7 Issue 1, 2015/11/1

に対する感受性)が神経伝達物質、とりわけグルタミン酸とセロトニンの機能に関わること¹²³、逆に脅威に対する感受性を弱め、新奇なものに対する感受性を強くする遺伝子が NMDA 型グルタミン酸受容体のひとつである NARG1 であること¹²⁴、それらと政治的指向性や音楽や絵画の好みなどと関連していること¹²⁵ などについても実証的な報告がなされている。

今日の近代社会に暮らす私たちは、隣の町の人間から襲撃される日常的な危険性にはさらされているわけではなく、「よそ者」に対する恐れや警戒心など、多くの場合は必要としていない。しかし、私たちの神経系オペレーティング・システムに何百万年もかけて組み込まれた恐怖モジュールは、外見や言葉や仕草などから「よそ者」を識別し、必要のない警戒心のスイッチを入れ、他者への偏見や嫌悪感を引き起こす。遺伝子に組み込まれた私たちのオペレーティング・システムが、それが全くの時代遅れとなった近代においても不必要に作動し、ときとして「よそ者」の非人間化のような顕在的行動をともなって、私たちの日常生活に深く根差しているのである¹²⁶。今日の日本でも、Twitter や Web で日本語の近隣民族に対する蔑称をキーワードに検索してみれば、同様の他民族に対する嫌悪感や攻撃性を簡単に見つけることができる。潜在連合テストから見出されたのは、そうした近隣の他民族に対する嫌悪感が、それを現実世界での意識的振舞いに露出する人たちだけのものではなく、私たちの大多数の日常的な意識の中に潜んでいるということである。そして、SNS 広告ネットワークの STEM 職広告、Facebook imSitu とマイクロソフト MS-COCO 、GloVe の 3 つの検証結果から見出されたのは、レベル 4 人工知能が、そうした人間の潜在的偏見を、人間の日常的な情景の画像やウェブ・クロールを通じて抽出し、増幅する特性をもっているということである。オックスフォード大学インターネット研究所のフェロー、サンドラ・ワッチャーが言うように、「世界には偏見が宿る。世界の過去データには偏見が内在する。人工知能が偏見を抱いたからといって別に驚くことはない。」¹²⁷

4.3.4 不得意分野はダイナミック・ヒューマニティ

このようにレベル 4 人工知能が掴み出す人間の潜在的偏見 (Implicit prejudice) は、人間の意識的思考 (Deliberate thought) レベルにおいて自覚される顕在的偏見 (Explicit prejudice) よりもかなり強いことも、いくつかの心理学的テストから明らかになっている。例えば、潜在的偏見と潜在的偏見の心理学的データを世界中から収集した Project Implicit¹²⁸の結果分析によると、

- 顕在的にアジア人に対して偏見をもっていると回答した被験者は、白人の被験者の 26.7%であったが、潜在連合テストでは、62.8%の白人被験者がアジア人に対する偏見を内在させている
- 同様に、年寄に対する偏見に関しては、若者の被験者のうちの 33.6%が顕在的偏見、78.9%が潜在的偏見を示している

¹²³ ジョナサン・ハイト 『社会はなぜ左と右にわかれるのか：対立を超えるための道徳心理学』 紀伊國屋書店 2014.04

¹²⁴ Pete K. Hatemi, Nathan A Gillespie, Andrew C. Heath, Nicholas G. Martin, “A Genome-Wide Analysis of Liberal and Conservative Political Attitudes”, *Journal of Politics* 73(1) · January 2011

¹²⁵ Kevin B. Smith et al, “Linking Genetics and Political Attitudes: Reconceptualizing Political Ideology”, *Political Psychology*, Vol. 32, No. 3, 2011

¹²⁶ Nick Haslam, *Dehumanization: An Integrative Review*, *Personality and Social Psychology Review*, 2006, Vol. 10, No. 3, 252–264

¹²⁷ Hannah Devlin “AI programs exhibit racial and gender biases, research reveals”, *The Guardian*, 13 Apr 2017

¹²⁸ <https://implicit.harvard.edu>

- 黒人に対する偏見に関しては、黒人以外の被験者の 34.0%が顕在的偏見、65.3%が潜在的偏見を示している

- 日本に対する偏見に関しては、日本人以外の被験者のうち 60.2%が顕在的偏見、78.7%が潜在的偏見を示している¹²⁹

いずれの場合でも、より多くの方が潜在的偏見をもっていた。

また、潜在的偏見は、顕在的偏見よりも行動に露呈しやすい。カナダで行われた模擬的な人事採用の心理学テストによると、アジア人に対する顕在的偏見と潜在的偏見が異なる 2 つのグループ、「顕在的偏見あり and 潜在的偏見なし」グループと「顕在的偏見なし and 潜在的偏見あり」を比較すると、実際の採用可否の意思決定行動において、より強いバイアスとなって露呈してしまうのは、「顕在的偏見なし and 潜在的偏見あり」の後者のグループであった。すなわち、意外にも、顕在的（また言語的）に自分が偏見をもっていることを自覚している被験者よりも、本人が自覚せず潜在的に偏見を抱えている被験者の方が、偏見が実際の行動に表出しやすいことが確かめられたということである。また、顕在的偏見も潜在的偏見も、「文句のつけようのない候補者」に対しては露出することがなく、採用可否を迷うような候補者に対して露出する頻度が高い傾向も、同じテストで観察されている¹³⁰。

日本経済新聞社は、「勘や経験をベースにしたアナログな判断」による人間の採用決定判断よりも、人工知能の方が、バイアスのない、公平な（日経新聞の言葉では「デジタルな」）判断ができると批評していたが、ここまでの人工知能の検証と心理学的テストのデータから、次のことが示された。

- 人間の顕在的な偏見量は、レベル 4 人工知能が人間の潜在的偏見から掘り出す偏見量よりも小さい
- レベル 4 人工知能は人間から掘り出された潜在的偏見をさらに増幅させる
- 人間の顕在的偏見は、あまり実際の行動には露出しない
- 人間の潜在的偏見は、（意識的思考によって）実際の行動に露出するのを抑制されている

したがって、日本経済新聞社が主張するような、「HR テック」による人間より「デジタル」で公平な判断など、レベル 4 人工知能に関する限りは到底できそうにもない。

一連の心理学的実験でさらに興味深いのは、潜在的偏見は非言語的な振舞いに表出し、顕在的偏見は言語的振舞いに表出することである。イギリス・ヨーク大学の心理学者、ケリー・カワカミらは、顕在的偏見と潜在的偏見が実際の人間の振る舞いにどのように露出するかを心理学的テストにより調べた。その結果によると、顕在的偏見は、言語的に振る舞いに若干影響して露出するが ($r=.40, p<.05$)、非言語的振舞いにはほとんど影響しない ($r=.02$)。一方、潜在的偏見は非言語的な振舞い（表情など）に影響して露出するが ($r=.41, p<.05$)、言語的振舞いには影響しない ($r=.04$)。また、顕在的偏見は、自分が

¹²⁹ Anthony G. Greenwald, Linda Hamilton Krieger, "Implicit Bias: Scientific Foundations", California Law Review Volume 94 Issue 4 Article 1, July 2006

¹³⁰ Leanne S. Son Hing, Greg A. Chung-Yan, Leah K. Hamilton and Mark P. Zanna, "A two-dimensional model that employs explicit and implicit attitudes to characterize prejudice", Journal of Personality and Social Psychology, 94, 971-987, June 2008

偏見をもっている相手に対して、嫌悪感を露出させることなくうまく振る舞えたという被験者の自己評価に若干影響する ($r=.33$, $p<.05$) が、潜在的偏見は、相手と第三者の評価に関連する (それぞれ $r=.40$, $r=.52$, いずれも $p<.05$) という¹³¹。

要するに、潜在的偏見 (Implicit prejudice) は顕在的偏見より強く、実際の差別的な行動に露出し、非言語的な振舞いに表出する。それとは反対に、顕在的偏見 (Explicit prejudice) は潜在的偏見よりは弱く、より抑制されており、実際の差別的な行動にあまり露出せず、言語的な振舞いに表出するのである。このことは、おそらく、社会心理学者や進化心理学者が「ソシオメーター」とよぶ心理的機能によって説明することができるであろう。ソシオメーターは、他人からどう思われているかを常にモニタリングし、集団からの排除が迫るほどに低くなると自意識に警報をならし、他人からの評価を改善する行動をとるよう設計されている心理装置である。人類は何百年もの間、小さな社会集団の中で協力して、他集団からの襲撃から身を守り、共同で食料を獲得し、子供を産み育てる適応課題に直面していたため、集団への帰属を認めてもらいか否かに自分の生存がかかっていた。そのため、自分が小さな社会集団の他の人からどう思われているかに極めて敏感になり、また、他人から好感をもたれるようにする能力が自然選択されてきた¹³²。

ソシオメーターは、恐怖モジュールと同じく、狩猟採集時代の人類の適応課題をこなすために備わった能力であるが、二つの点で、恐怖モジュールとは大きく異なる。ひとつは、恐怖モジュールの働きが今日の社会での生存にとっては時代遅れの能力となっているのに対し、ソシオメーターは今日の複雑な社会においても有効な能力であるという点である。SNS は他者評価に敏感なこの人間心理ソシオメーターを充足するツールであり、それに支えられて成立しているといってもよいし、社会一般に目を向けると、法律に規制されていなくても、社会的評判を気にするほとんどの人が様々な社会的ルールを守る「相互評価システム」によって安定的な社会が成立している。もうひとつ大きく異なる点は、恐怖モジュールが、人間の心的システムの中にある多数のモジュールの内のひとつであり、遺伝的要因、環境要因によりスイッチが入りやすい傾向の人と入りにくい傾向の人がいて、どちらのタイプもそれぞれ自然選択されてきたと考えられている¹³³のに対して、ソシオメーターは、遺伝的特質に関わらず人間が普遍的に備え、環境に関わらず常にスイッチがオンになっているという、人間行動に基盤的な役割を果たしてきたとみられている¹³⁴という点である。

ソシオメーターは、心理学的テストにおいても、他者からの肯定的評価は自尊心などの感情に対して、他者からの否定的評価は孤独感・憂鬱・嫉妬などの感情に対して、大きく影響することが確かめられてい

¹³¹ John F. Dovidio, Kerry Kawakami, Samuel L. Gaertner, "Implicit and Explicit Prejudice and Interracial Interaction", JAN 2002, Journal of Personality and Social Psychology, VOL.82, ISSUE.1, 62–68.

¹³² Lee A. Kirkpatrick, & Bruce J. Ellis, (2001). An evolutionary-psychological approach to self-esteem: multiple domains and multiple functions. In G. J. O. Fletcher & M. S. Clark (Eds.), Blackwell handbook of social psychology: Interpersonal processes (pp. 411-436). Oxford, UK: Blackwell Publishers.

¹³³ ジョナサン・ハイト 『社会はなぜ左と右にわかれるのか：対立を超えるための道徳心理学』 紀伊國屋書店 2014.04

¹³⁴ Lee A. Kirkpatrick, & Bruce J. Ellis, (2001). An evolutionary-psychological approach to self-esteem: multiple domains and multiple functions. In G. J. O. Fletcher & M. S. Clark (Eds.), Blackwell handbook of social psychology: Interpersonal processes (pp. 411-436). Oxford, UK: Blackwell Publishers.

る¹³⁵。また、偉大な業績を残した 91 人の芸術家や科学者を対象に行なったインタビューから創造性を育む要因を調べた心理学者のミハイ・チクセントミハイは、創造性を育むその重要な要素として、年長者や指導者からの「正当な評価」、および、家族内で始まり、仲間集団、学校、一般社会へと継続される「期待」という二つの（いずれも他者からの評価に関わる）精神的な支援をあげている¹³⁶。また、行動経済学者のジョージ・レーベンシュタインらは、いくつかの他の心理学的テストをふまえつつ、次のように述べている。「（人間の）意識的思考システムが持続的に関心を抱く他者は、人はどのように振る舞うべきかについて道徳的倫理的原理によってつき動かされる他者である。対照的に、（人間の）情動システムは、引き金となる共感の程度に応じて、純粋な利己心と極端な利他主義の間のどこかに向けてつき動かされる」¹³⁷。

つまり、この他者の評価に敏感な「ソシオメーター」が、「差別主義者」などと他者から否定的な評価をされることを回避するように機能し、それが潜在的偏見を顕在化させることを軽減し、他者とのコミュニケーションにおいて言語的に表出してしまうことを抑制し、雇用差別的な行動となって露出することを抑制していると考え得るのである。レベル 4 人工知能が、ソシオメーターが抑制している潜在的偏見を掘み出し増幅させてしまう原理的習性をもっていることは、Facebook imSitu vSRL、Microsoft COCO MLC、GloVe の 3 事例から確かめられた。もう一方で、この潜在的偏見を抑制しているソシオメーター自体に対しては、レベル 4 人工知能は、どのように振る舞うのだろうか。潜在的偏見を抑制し、創造性を喚起し、偉大な芸術家や科学者を育て、道徳的倫理的原理によってつき動かされる他者に持続的な関心を抱き続けるような人間の社会性を、レベル 4 人工知能は特徴量として掘み出せるのだろうか。

レベルの 4 人工知能の原理的構造を考慮する限り、どうもそんなことには適していないようにみえる。第一に、「制御性：人工物理システム（ハード 2 類）」の項で述べたように、レベル 4 人工知能は大量のデータへの依存性が高く、データの「癖」と抽象的な振舞いの不変性をほとんど切り離せない。人工物理システム（ハード 2 類）の場合は、環境 X を極めて小さな範囲に限定することで、データの「癖」が強くなり過ぎて著しく不正な振る舞いにつながらないように施されていた。しかし、「HR テック」の場合、機能対象（環境 X に相当）は、「社員の人物像や特性」「働きぶり」であり、出力（振る舞い Y に相当）は「評価や最適な職場」「配置転換」「採用可否」であるので、「人物像」や「働きぶり」を特定の狭い範囲に限定してしまうと、その限定自体がバイアスとして働く。したがって、公平な評価や配置転換や採用可否の判断は出力できないということになる。仮に、もしも、「人物像」や「働きぶり」を網羅的にレベル 4 人工知能に学習させるだけの大量の指数関数的に増量されたデータを集め、アノテーションに工夫をこらすことができた（と開発者が思った）としても、それが、「人物像」や「働きぶり」に対して網羅的であり、公平な「評価や最適な職場」「配置転換」「採用可否」の判断を出力しているかを検証することは、IEEE も言う通り、「ブラックボックス」のレベル 4 人工知能には「予見可能な未来において、

¹³⁵ Mark R. Leary, Ellen S. Tambor, Sonja K. Terdal, Deborah L. Downs, “Self-Esteem as an Interpersonal Monitor: The Sociometer Hypothesis”, *Journal of Personality and Social Psychology* 1995. Vol. 68. No. 3. 518-530

¹³⁶ ミハイ・チクセントミハイ 『クリエイティビティ：フロー体験と創造性の心理学』 世界思想社 2016.10

¹³⁷ George Loewenstein, Ted O’Donoghue, Sudeep Bhatia, “Modeling the Interplay Between Affect and Deliberation”, *American Psychological Association* 2015, Vol. 2, No. 2, 55-81

おおよそ有り得そうにもない」。公平性が検証できなければ、公平とは認識できないので、やはり、公平な判断は下せないことになる。

第二に、レベル4人工知能の一般化能力の問題がある。2018年5月にバンクーバーで開催された ICLR 2018 では、再帰型ニューラル・ネットワーク (Recurrent neural network : RNN) の一般化能力の問題が報告されている¹³⁸。その報告によると、学習データに含まれていたことがないパターンの新しい文章を翻訳するタスク (例えば、「左と反対に歩いて2回走る」を翻訳させる) においては、6-7割程度の高い精度を出せた。しかし、新しい動詞を学習させて単純な文章を作る (例えば、「スタ爆」(「スタンプの連打」という意味の SNS 言葉) という新しい単語の翻訳を学習させて、「スタ爆してください」といった簡単な文章を翻訳させる¹³⁹) というシステムティックな一般化能力を必要とする翻訳タスクにおいては、1-2割程度の低い精度しか出せなかったという¹⁴⁰。Google の人工知能研究者で Keras 開発者のフランシス・ショレットが指摘するように、人間には、「既知の概念と一度も経験がないものの表す概念を融合する「極度の一般化能力」(extreme generalization) がある。これは、レベル4人工知能の動作とは著しい対照をなす。それは「『局所的な一般化』(local generalization) とでも呼べるもので、ディープ・ニューラル・ネットワークが実行する入力から出力へのマッピングは、新しい入力が訓練時に経験したものと少しでも異なる場合、即座に意味をなさなくなる」¹⁴¹。認知心理学者のスティーブン・ピンカーは、そういった人工知能の開発者らによる「極度の一般化能力」は、それでもまだ人間の能力を過小評価していると批判したうえで、人間の一般化能力を「解釈による一般化」とし、次のように定義づけている。人間には、「対象物を心的に<シンボル化>し、それらの心的シンボルを、同じく心的な複数の規則系に照らして操作することができる『解釈による一般化』能力があり、知識を数量化した再帰的命題の形にし、相互に関連しあって、経験のある特定分野についての『直観的理論』を形成する」¹⁴²。

ピンカーのいう人間の「直観的理論」は人事の現場で威力を発揮する。例えば、ある応募者との面接に際して、このスタッフは、標準的な基準に照らせ合わせれば、それほど高く評価できるわけではないが、そこからみ出る「何かをもっている」と直観的に閃き、それに「かけてみよう」と考える。あるいは、このスタッフは遊ばせてみよう、その方がいい仕事をするに違いない。そういう直観的な判断が企業人事の現場では時々発生し、だいたいにおいて、意外に優れた成果を生み出すのはそういう場合であったりもする。個別の人間に関わるレベルだけではなく、組織的レベルでの人事的判断も、直観的理論に支えられている。例えば、この組織には迅速性を醸成する必要があると判断し、組織風土改革プログラム「ガゼル」を起ち上げようとする。あるいは、この組織では社内政治が蔓延していると感じて、次年度の業

¹³⁸ Brenden Lake, Marco Baroni, "Still not systematic after all these years: On the compositional skills of sequence-to-sequence recurrent networks", ICLR 2018 : International Conference on Learning Representations, Vancouver, Canada, Apr 30, 2018 - May 3, 2018

¹³⁹ 実際の検証で使われたタスクは、「dax」という新しい言葉を使って、「he is daxy」などの文章を翻訳させるタスク

¹⁴⁰ Brenden Lake, Marco Baroni, "Still not systematic after all these years: On the compositional skills of sequence-to-sequence recurrent networks", ICLR 2018 : International Conference on Learning Representations, Vancouver, Canada, Apr 30, 2018 - May 3, 2018

¹⁴¹ Francois Chollet "The limitations of deep learning", 17 July 2017, In Essays adapted from Section 2 of Chapter 9 of Francois Chollet "Deep Learning with Python" (Manning Publications) <https://blog.keras.io/the-limitations-of-deep-learning.html>

¹⁴² スティーブン・ピンカー『心の仕組み：人間関係にどう関わるか(上)』日本放送出版協会 2003年06月

績評価は、顧客中心主義と社内政治撲滅を柱にしようとする。それらをスタッフ・フォーラムのスピーチに、格調高い人文学的レトリックに織り込もうと考える。それらは、人事コンサルティング・ファームの出すデータや提案（概ねレベル 2-3 人工知能で構成されたデータであるが、レベル 4 の分析結果も興味深い）を参考にはするものの、全体的・総合的には指導者の直観的理論によるところが大きい。著名な投資家ウォーレン・バフェットの名言「あなたが一生を添い遂げてほしいと思うような企業に投資しなさい」は、直観的理論による投資意思決定を推奨したものであるが、直観的理論の投資家に選ばれたいと願う企業経営者は自ずから直観的理論に優れた経営者であろうとし、それを外に向けてアピールもする。レベル 4 人工知能の「類似性による一般化」¹⁴³においては、そうした人間の直観的理論は、「解釈による一般化」に照らされることがないために、それが優れたもの卓越したものであるほど、ビッグデータの平準的類似性の中にうずもれ、抽出されにくくなる。

第三に、レベル 4 人工知能の特徴量抽出に貫かれているロジックは、類似性であって、因果関係ではない。心理学者が開発した性格診断や適性診断モデルは、レベル 2-3 人工知能で構成された「人物像」の一端を説明するツールであるが、質問の回答と診断結果の関係を明確に論理的に説明することができる。また、レベル 1-2 人工知能で集計される企業の 360 度業績評価は、「働きぶり」の評価ツールであるが、どの企業でも統一的な評価基準が明確に示され、各評価者はなぜその評価をつけたかを明確に論理的に説明しており、それが被評価者にも透明にされている。レベル 1-3 人工知能で開発されたこれらの人事ツールを、すべてレベル 4 人工知能により置換した場合を想定すると、次のようなことになる。まず、レベル 4 人工知能が出力した人事評価は、同じような振舞いをした人と同じような特徴量がマッピングされているにすぎないので、結果について被評価者がわかるのは、自分は誰と似ているのかという類似性であって、自分がどんな振る舞いをしたからこの評価が下されたのかの因果関係（そのスタッフの来年度の改善や奮発の動機づけになる）はわからない。また、その類似性は、付与された学習データの範囲に限定され、かつ、その学習データからレベル 4 人工知能が学習できる範囲（例えば「局所的な一般化」の範囲）に限定されるため、例えば、業績に各段の差がある Aさんと Bさんが、同じような人物像や働きぶりの評価を下されるということが理論的には起こり得る。そして、その場合でも、どの範囲での類似性なのかという因果関係は、被評価者だけではなく、開発者にもわからない。（このように入力と出力の因果関係がみえないという意味では、レベル 4 人工知能による人事的判断は、星座占い性格診断に類似するのであるが、星座占い性格診断の場合、ひとつのベクトルがひとつの生得的属性に均等に割り当てられる（誕生日による星座の割り当て）という点で公平であり、また、誰でも自分自身の経験から、この性格診断は当たっているとか、外れているとかといった判断ができるので、透明性という意味でも公平である。）

第四に、類似性原理に貫かれて特徴量をマッピングしていくレベル 4 人工知能を通してみえる世界は、フラットで静的な世界である。そこでは、論理的定量化も、抽象化も、構造化もされない。すべての特徴が対等に、あたかも長いリストのようにフラットに位置づけられ、他のレベル 4 人工知能のプログラムから学ぶことも難しく（「再利用性」問題）、その学習データからセンサリングできる範囲において閉じられ、その学習データが集められた時点において静的なスナップショットである。一方、現実世界の人間

¹⁴³ スティーブン・ピンカー『心の仕組み：人間関係にどう関わるか（上）』 日本放送出版協会 2003年06月

は、きれいにパッケージ化されたデータを定点観測で都合よく提供してくれるわけではない。他者からの評価に敏感なソシオメーターに基づき振る舞う人間は、他者との対話によって勇気づけられ、発奮・鼓舞され、共鳴し、共感し、考えを改める。他者との関わりと通じて、創発的に、ダイナミックに、散発的、偶発的、揮発的に振る舞う。採用面接の現場において、面接官の指導者の心的システムの中心にあるのは、箇条書きの長いチェックリストなどではなく、候補者との対話から創発的なダイナミズムを生み出そうとする欲求であり（そして、そのような創発的発火が起これば採用、起こらなければ不採用にする場合が多い）、創造性やイノベーションを生み出す源泉は、類似性から静的に導き出される特徴量によって示される能力などではなく、職場の上司や学校の指導者からの精神的支援による創発的動機づけである。

この人間の創発的・創造的な動的な世界について、コレージュ・ド・フランスの歴史学者、マルク・フュマロリは次のように言い表している。「ヒューマニティについて語る時、私たちは、遠い過去の最高の詩人、最高の芸術家たちの驚くべき多様性と新奇性が結実した最高の〈ところ〉と対話しているのである。それは、世代から世代へと驚くべき受精を通じて受け継がれ、新たな果実を見い出され、内面化されていく」¹⁴⁴。そうした〈ところ〉の対話に伏流するのは、西欧的文脈では、ルネサンスの人文主義に始まり、啓蒙主義や実存哲学、近代の様々な新しい価値観が創造され、近代社会へとつながっていくダイナミズムであり、日本の文脈では、院政期から室町時代にかけて始まった「一芸に秀でる」ことを有徳とする求道主義。それを実践する表現する職能民たちの精神的世界を支えた『至歌道』『徒然草』などの文学を通じて、革新的な商工業文化が醸成されていくダイナミズム。それが幕末の通商条約締結を契機に近代商工業化へと爆発するダイナミズムである。企業組織の現場においても、率いる組織の規模が大きくなるほど、指導者にはこうしたヒューマニティ的素養が求められる。しかし、こうした人間の創発性のダイナミズムは、レベル 4 人工知能の類似性原理のマッピングを通してしまうと、付与された学習データの狭い偏った世界に閉じこめられた、フラットで静的で平準的な特徴量に変換されてしまうのである。

公平な判断を出力することができず、人間の創造性の源泉となる直観的理論や他者との創発性のダイナミクスを取り込むことができない。そんなレベル 4 人工知能が、「HR テック」のような人間活動システム（ソフト 3 類）に適しているなどと想像するのはとても難しい。

¹⁴⁴ Marc Fumaroli, "Les humanités au péril d'un monde numérique 2/2", Le Figaro, 31/03/2015

5 次世代人工知能を構想するためのアプローチ

ディープラーニングの限界を確認した後に取るべき方法として二つの道がある。第1には、ディープラーニングの方法が有効な範囲でそれを活用し、また有効性が保証されない領域での使用結果に対しては懐疑の目を向けその不当性や過誤の原因を明らかにして誤用を回避する。たとえば「AI で強化された自動トレードシステム」等の触れ込みを煩瑣に見かけるが、その内容を分析することで誤用のメカニズムを明らかにすることが出来る。第2には、ディープラーニングを初めとする機械学習の方式では扱うべきでない広大な領域に対してどのように立ち向かうかという、本来の意味での本格的な「次世代 AI」の開発という新たな課題がある。本報告書では、現在急速に普及が図られている第1の道ではなく、人工知能活用の実態分析を通して得たヒントから、以下に第2の道「次世代 AI の構想」のためのアプローチについて最後にまとめたい。大きく区分すると3種のアプローチがある。

5.1 哲学的アプローチ—認識論と知識論の精査

本報告書では科学主義の哲学の立場をとった¹⁴⁵。ここでは「確かな知識」の根拠を「外部」に求める。「内部」を根拠とする「確信」等は移ろい行くものとして知識のアンカーには位置づけない。その場合、「ハード2」（意思的人間・集団）に対する「ソフト2」は常に近似的モデルに過ぎない。またその動態である人間活動システムに対する「ソフト3」に対しては、通常、モデルによる近似すらおぼつかない。事例分析で見たように厳密性を旨とする場合、この領域に対してはディープラーニングの適用を避けていた。

ところで、科学主義の哲学のみがヒトの認識現象を支配しているわけではない。知識の科学的正確さを求めるならば科学主義の哲学に依存せざるをえない。しかし、ヒトはたとえば「確信」に基づく意思の伝達や意思決定も日常的に行っている。そこで、認識論や知識論の新たな基盤となりうる哲学的アプローチとして、「確かな確信」を求める立場を追究してみてもどうかであろうか。たとえば現象論の哲学や実存の哲学である¹⁴⁶。ここでは反復的対話を通して「確かさ」が互いに確認されたことにより「確信」が伝達されたとする。したがってAIの内容は個固有の価値観を反映したものとなるであろう。

5.2 事例分析によるアプローチ

本報告書では、日本経済新聞、人工知能学会誌、プリファード・ネットワークス社の実績、国家プロジェクトとNEDOで実施した研究開発の4種類の事例群を取り上げ、その内容を科学主義的知識論の原理により分析し、現在各所で進められている人工知能の実態の把握に努めた。「次世代 AI の構想」に際しても、その具体的な姿を想定するためには同様に事例分析を欠かすわけにはいかない。ただし、その事例は現在限られていて多くはない。

第1には「人文学」をとりあげるべきであろう。科学主義の枠組みで人文現象を考察する「人文科学」とは異なる分野である。

そして第2には先進的な映像の世界であろう。たとえば、仮想現実や拡張現実を組み込んだアニメの世界やプロジェクションマッピング等である。今夏、七夕に合わせて製作された「約束の七夜祭り」は

¹⁴⁵ 戸田山和久 知識の哲学 産業図書 (2002)

¹⁴⁶ 小林隆児、他 人間科学におけるエビデンスとは何か 新曜社 (2015)
植村玄輝、他 現代現象学 新曜社 (2017)

YouTube で配信されていて、その構成を分析すると以下のようになっている。

「約束の七夜祭り」AI 要素（その他、AR・VR 技術とその応用）の分析

【ストーリー】主人公が友人から呼び出されて向かった山奥で道に迷ってたどりついた場所は、不思議なお祭りを行う村だった。

制作：XFLAG スタジオ

公開日：2018 年 8 月 3 日より

配信サイト URL：https://nanayamatsuri.com、youtube

要素の抽出とその評価内容

時間	内容等
2 : 5 4	この世界の人々が使う「トータルライフサポートアプリ」。さまざまな機能が無料で使える。主人公は使用している友人に勧められて子供の頃にインストールした。（これにより、様々なライフログが収集されていた。）
5 : 3 0	ホログラムによる仮想人格の表示
1 0 : 2 0	スーツ（憑衣、烏帽子、モーション取り込み用&反応のフィードバック用）
1 3 : 0 0	ホログラムが集落の中を自由に動き回る・・・、これは、どのようにすれば可能か？
1 6 : 4 5	HMD を使った VR で作業手順指示をしてくれる、服にパワーアシストがある服が生体情報を取得してサーバに送信してくれる、パワーアシストだけではなく筋肉へ指示を出すことで、作業したことが無くても指示通りの動きが可能になっている → VR での指示は NEC などが実用化している ただし、このアニメの世界では具体的な人間の作業により物理的な建物が作られるのではなく、あくまで仮想空間内の建物をプログラミングする（ホログラムのもとを作り出す）に過ぎない。この過程が通常のプログラミング作業としてではなく、あたかも通常の人間の建設作業として描かれているのである。 → VR・AR で作業すること＝プログラミング作業の未来像。 昨今の子供向けのプログラミング教育もコードを描かせるようなことはしない。（誰にでもできるようにするため。）
2 0 : 3 0	仮想の敵に刀で切られた → スーツへの電気信号レベルで実現可能？ 敵が登場するがそれもすべてホログラム？

23 : 45 祭りがコンピュータに仕切られていて、参加者はコンピュータに招待される

38 : 00 かつて存在した仮想人格（武者）が暴走する。もともとこのアプリは「ライフログアプリ”ほしのめぐり”」で、そこで収集した様々な人のライフログを統合してユーザーの人格の複製を仮想人格としてバーチャル空間内に生成する。そして、実在の本人と仮想人格のズレを検証し修正を繰り返すことでその精度をあげてきた。その過程の中で、ライフログが存在しない部分につき、文章や絵の中にしか記録が残っていない過去の人物の仮想人格を蓄積した膨大なパターンから近似的に構築する技術を学んだ。（人格のコピーではなくねつ造と言っている）← 強いAI

仮想人格が仮想と現実の世界をつなぐ扉のすき間をすり抜けて出てきた。（プログラムのバグによる暴走？）

「カンナ」は仮想人格だが、自身を生み出すソフトウェアの修正もできる。（ソフトウェアによるソフトウェアの自己修復、自己プログラミング）

40 : 20 武者たちを倒す武器が主人公たちに与えられるが、これは武者たちを消去するプログラムであり、スーツに入れて（憑衣にスキルを降ろしたという表現）使うことができる。（ポケモンGOにおいてAR上でボールを投げるようなイメージ）

41 : 00 建物の床はホログラム（多分、舞台自体がそのような感じになっている、4D?・・・）

41 : 42 「父上と家臣たちの無念わらわがはらしてくれよう」といいながら、敵キャラが切りかかってくる（消去されようとしているプログラム（?）がそれを阻止しようとするような行動を示す反応）

仮想現実内の物語の中でキャラクター（NPC：ノンプレイヤーキャラクター）が動くイメージは、昔からあまり変わっていない（マトリックス、ネットフォース）。

ただし、単にファイアウォールなどのような侵入者を防ぐ機能を擬人化して見せるだけなのか、相手の行動に合わせてこちらの意思を示すような行動を取る（AI）のような違いはあるが・・・（AI学会誌、ゲームとAI他にもこのテーマは登場する。）

また、プログラムなどを擬人化して、仮想的に人間に見せる仕組みはデジタルゲーム自体がこの仕組みになっている。

45 : 43 この祭りを完遂するのがプログラム（カンナ）の役目なので、祭りが終わったら消えてしまった。（＝プログラムの実行が止まる）主人公たちは、擬人化されたプログラムを人に接するように感じている。

47 : 45 主人公の友人は既に亡くなっていて、仮想人格がメールや電話を主人公に送ってきていた

50:00 確かに、ここに出てくる仮想人格の人々には「心」はないだろう・・・

死者との再会という設定ではあるが・・・

プロフィールされた情報から、「こういう状況だったら、このように振る舞うだろう」ということを作り出しているだけである。もしかすると、本人も区別できないくらい同じ反応作り出せるかもしれない。(心の中をのぞくのは、自分自身であってもできないため、作り出された反応を「あなたはこう反応しました」と見せられたら、そうか・・・とってしまうかもしれない。)しかし、本人にしか本当に知らないこと(ライフログとして収集できない情報)は、作り出せないため、100%再生は無理なのか?

54:30 既に死んでしまっていた友人と接する主人公。反応は友人そのものだけれども、普通にそう反応するだろうなというレベル?

受け取る側が本当に友人と相対していると思えば、そう思える世界観が表現されている。

受け手が人間なのでそれを許容できる。

死者に会えるという話はたくさんあるが、膨大なライフログデータに基づき亡くなった人や過去の人々を仮想人格として生成し、それと対話することで仮想的に死者と会う(会話できる)ことができる世界は実現可能だと感じる。死後の世界があって、死者がそこからこの世にやってくるというような話よりは現実味がある。

この作品は、来年度概算要求に載せられた「ムーンショット型研究開発制度」の事例を先取りしているとも言える。ムーンショットのプロジェクト事例に「仮想現実の中で生き返らせた故人と会話ができる技術」とある。その実態的姿が、既に映像を伴う物語として提示されているのである。

5.3 脳科学や心理学からのアプローチ

脳科学や心理学の領域は人工知能の実体的側面を担っている¹⁴⁷。人工知能で実現し難い課題として、自我とか意識の世界がある。個体としてのヒトの自我や意識の多くは個人的な経験の集積により形成されるといわれているが、それだけならばスマホ様の記録装置にライフログと心情の変化を記録しておけば形成できそうにも思えるが、一方で生物としてのヒトには長い進化の歴史があり、その過程で先見的に作りこまれた実体が脳として存在している。この先見的記憶は脳の進化過程と当時の体験や環境条件により刷り込まれたものであろう。

以上に示したような多面的なアプローチを深め総合化することにより、次世代の人工知能を構想することが有効であろう。

¹⁴⁷ 松田雄馬 生命から紐解く知能の謎：人工知能の哲学 東海大学出版部 (2017)