システム監査学会2020年度　第34回研究大会

「AIと個人情報～AIの諸問題と個人情報保護～」プロジェクト ＆個人情報保護専門監査人部会 報告 「AI及び個人情報とシステム監査について」

2020年６月5日

報告者　朝倉 俊道 CISA

1

目次

2

経緯

現在は社会・経済・ビジネス・生活のあらゆる面において、AI（Artificial Intelligence）が謳われています。ベースとなりますデータ、ビッグデータを分析・解析し、これからのビジネスを動かし、人間生活を支えるのがAIです。また、AIはDX（デジタルトランスフォーメーション）の重要な要となります。

AIにはデータが必要です。街頭にカメラを設置して無断で録画した情報（肖像権）、インターネットにあるコンテンツをスクレイピングツールにより抽出した情報（著作権）等、個人の情報であり個人の権利です。即ち、「個人情報／個人の権利」は、「AI」に常につきまとう課題です。そして、プログラムにより実装されるシステムと同様に、AIを用いたシステムもPL（製造物責任）から逃れることは不可です。加えて、学習に用いるデータは、従来のデータを消費する分析とは異なり、形を変えて永続します。このことにより、商用システムにおけるAIの使用は、従来のアルゴリズム実装よりも様々なリスクを抱えています。

また、2017年5月に、約10年ぶりに改正された「個人情報保護法」の全面施行、2018年5月に施行されましたEU の一般データ保護規則（General Data Protection Regulation：GDPR）、2017 年5月に120年ぶりに国会で可決された改正民法が2020年4月に施行されます。ITシステム・サービス等の業務委託契約に関連するところでは、「瑕疵担保責任」が、契約内容に適合しない場合に修補・追完を請求できる「契約不適合責任」に変更されました。これにより、契約時における契約内容の明確化がより一層求められるようになりAI活用におけるデータの透明性、責任、知的財産、契約等、種々なリスクと対応すべき課題があります。

このように想定されるAIのリスクを適切にコントロール・運用するためシステム監査はどうあるべきかについて調査・研究を行っています。今回、中間報告です。

3

1. 活動メンバー

個人情報保護専門監査人部会メンバ
研究プロジェクトメンバ　6名

氏名	所属		個人	研P
朝倉　俊道	エムビーケーメタルソリューション 株式会社			●
稲垣　隆一	稲垣隆一 法律事務所	主査	●	
黒澤　兵夫	TAKE国際技術士研究所	副主査	●	●
東野　憲康	株式会社 ムーンライトシステム		●	●
牧野　博文	株式会社 東芝		●	●
芳仲　宏	「法とシステム監査研究」プロジェクト		●	●

4

2. 活動状況

- 月に一回ペース　平日　18:30開始　約2時間

回数	開催日	場所	内容
第1回	9月6日	稲垣法律事務所	今後の進め方等
第2回	10月9日	稲垣法律事務所	資料の検討と報告
第3回	11月26日	稲垣法律事務所	「おそろしいビッグデータ」検討等
第4回	1月14日	稲垣法律事務所	「AI・ロボットの法律実務Q&A」検討等
第5回	2月10日	稲垣法律事務所	研究発表大会内容について（AI資料の検討等）
第6回	3月26日	稲垣法律事務所	研究発表大会内容について（課題等）
第7回	4月中	メールでやりとり	まとめ

5

3. 資料／文献調査

参考文献

No	書名	著者	出版社	発行年月日
1	おそろしいビッグデータ 超類型化AI社会のリスク	山本 龍彦	朝日新書	2017年11月30日
2	AI・ロボットの法律実務Q&A	第二東京弁護士会	勁草書房	2019年2月20日
3	機械学習＆ディープラーニングのしくみと技術がこれ1冊でしっかりわかる教科書	(株)アイデミー 山口 達輝／ 松田 洋之	技術評論社	2019年9月14日
4	AIの時代と法	小塚 荘一郎	岩波書店	2019年11月20日
5	AI・データの利用に関する 契約ガイドライン 1.1版	経済産業省		2019年12月

6

3. 資料／文献調査（1-1）

（1）「おそろしいビッグデータ　超類型化AI社会のリスク」山本 龍彦

【概要】

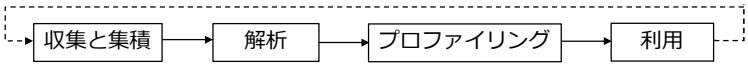
・ビッグデータやAIはそれ自体「悪」ではない。本当におそれなければならないのは、ビッグデータに基づくAIの評価が個人に関する100%正しい評価として自動的に受け入れられてしまうこと、憲法の基本的な価値を否定するような方向で使われてしまうこと、である。

・個人情報の漏えいよりも恐ろしい、第三者による「プロファイリング(個人分析)」がもたらす人権侵害や憲法問題を、憲法論の観点から分かり易く説明。

【選択理由】ビッグデータやAIを利活用していく上での「おそろしさ」を理解し、ビッグデータやAIの適切な利活用のあり方を考える上での種本としたいという思いから。

【トピック】　※以降全て【出所】「おそろしいビッグデータ 超類型化AI社会のリスク」山本 龍彦

①ビッグデータ利用の基本的なサイクル



②ビッグデータ社会におけるプライバシー権の限界

（a）本人の「同意」が怪しい

（b）普通の個人情報のコントロールが不十分

（c）匿名データの集積プロセスの中で突然個人特定性が現れるケースへの対応が不十分

③AIプロファイリングにもエラーやバイアスが紛れ込む

⇒採用や融資といった場面で不利益を受ける可能性有り

④個人化の嘘

ビッグデータに基づくAIの予測・評価は、「あなた」個人ではなく、グルーピングされ類型化した「あなた」に対するもの

⇒人事採用、融資、保健、教育といった場面で、その評価に反論出来なくなる

7

3. 資料／文献調査（1-2）

⑤神格化されるアルゴリズム

自動化バイアス＋AIの評価・判断に対して何が間違っているか具体的に説明できない事

⇒AIの評価・判断に反論できなくなる

⑥憲法の「個人の尊重」原理に抵触する問題

（a）「人生をやり直す自由」の侵害

（b）「血」（遺伝的事情）に拘束される個人

（c）過去の差別の助長

（d）誰もが対象になりうる新たな差別

⑦個人情報保護法の死文化

事業者は、フツウの個人情報を集めてプロファイリングすることで、本人の同意を得ることなく要配慮個人情報（センシティブ情報）を「入手」可能

⇒日本では「情報漏えい」問題が焦点で、プロファイリング問題に対策が講じられない

⑧GDPR（一般データ保護規則）の先端的な規定

（a）プロファイリングに対して異議を唱える権利（中止請求権）

（b）自動処理のみに基づいて重要な決定を下されない権利

（c）透明性の要請

⑨憲法論の必要性

・EUやアメリカでは、「ビッグデータのおそろしさ」が、憲法レベルで議論され、法的にもそれなりの対応策がとられてきている。

・日本は立ち遅れており、ビッグデータやAIの利活用に関する日本の議論は、経済合理性や効率性ベースで進んでおり、憲法上のリスクを語ろうものなら、「空気読めよ」という感じで白い目で見られる。

8

3. 資料／文献調査（1-3）

⑩自己情報コントロール権を鍛える

「コントロール」対象の拡大

- (a)プロファイリングの普及を見据えて、自分の管理下にないところで「事実らしく受け取られる情報」(とりわけセンシティブな情報)が生み出されること(推知されること)に対しても、本人のコントロールを及ぼしていく
- (b)データベース上に記録された自らの「過去」に対するコントロールを強化していく
- (c)自らが「つながる」ネットワークシステムを主体的に選択する権利を強化していく
- (d)個人が情報管理の構造やシステムに対して一定のコントロールを及ぼしていく

「コントロール」の実効性の強化

- (a)情報の利用のあり方や情報のフローを、個人が直感的に理解できる告知の方法を要求する
- (b)サービスを「人質」にとって同意以外の選択肢を事実上奪うことを禁止する

⑪本当におそれなければならないもの

・ビッグデータに基づくAIの評価が、「あなた」に関する100%正しい評価として自動的に受け入れられてしまうこと、あるいは、民主主義を歪めるようなかたちで利用されてしまうこと、つまり、憲法の基本的な価値を否定するような方向で使われてしまうこと

⇒ビッグデータが憲法と調和的に利活用されているかどうかを、常に監視していく 9

3. 資料／文献調査（2-1）

（2）「AI・ロボットの法律実務Q&A」第二東京弁護士会

【概要】

・急速に進展するAI・ロボットの社会的なインパクトに対して法律実務も対応していく必要があることから、AIのような新しい問題に対して、法学分野の見地から問題点を洗い出し、実務状況に基づいて、留意すべき点、問題への対応を検討している。

・構成としては、憲法分野、民事法分野、刑事法分野、行政法分野、知的財産法分野、国際問題のテーマに分け、それぞれの分野におけるAI・ロボットの法律問題をQ＆A方式でコンパクトに解説している。体系的なものではなく、注目されている論点に絞っている。

【選択理由】AIを利活用していく上で起きてくる法律問題に目を通すことで、AIの法律面からのリスクおよび個人情報保護に関わる問題に対する理解を深め、整理したいという思いから。

【トピック】

①AIの概念 **※以降は全て【出所】「AI・ロボットの法律実務Q&A」第二東京弁護士会**

・**AIは、推論・判断などの人間の知的機能を人工的に実現するための研究またはこれらの機能を備えたコンピュータ・システム**である。

・AIは、人間の知能そのものをもつ機械を作ろうとする立場からの汎用的なAI（**「強いAI」**）と、人間が知能を使ってすることを機械にさせようとする立場からのAI（**「弱いAI」**）に大別でき、主に実用化が進められているAI技術は機械学習を用いた弱いAIで、現時点でのわが国の契約法務でも弱いAIを念頭に置いた対応がなされている。

②（憲法）家庭用ロボットやAIを使ったプロファイリングには、プライバシーとの関係でどのような問題があるか？

・家庭用ロボットにより、私的領域として守られてきた生活空間においてプライベートな情報が収集され、これまで想定されなかった新たなプライバシー侵害が生じるおそれがある。

・AIを使ったプロファイリングでビッグデータを分析することにより、個人の私的事項が推知され、その人の「事実・真実」として取り扱われ、プライバシー侵害のおそれがある。

・**プライバシー・バイ・デザイン**の実践により、実効的なプライバシー保護が実現する。10

3. 資料／文献調査（2-2）

③（民事）**AIを利用したソフトウェア開発を委託する契約**を締結するにあたり、どのようなことに気を付ける必要があるか？

・経済産業省が2018年6月**「AI・データの利用に関する契約ガイドライン」**を作成・公表

AIを利用したソフトウェア（特に学習済みモデル）の開発の特徴

（a）学習済みモデルの内容・性能等が契約締結時に不明瞭な場合が多い
（b）学習済みモデルの内容・性能等が学習用データセットによって左右される
（c）ノウハウの重要性が特に高い
（d）生成物について更なる再利用の需要が存在する

↓ 契約時に先を見通すことが困難

「探索的段階型」の開発方式を提唱

①アセスメント段階 → **②PoC段階** → **③開発段階** → **④追加学習段階**

【ユーザ目線】**開発を含めた全体のコスト把握が困難**となるデメリットがあるが、途中の段階に進まないという選択ができる点では、双方にとって無用な損害の拡大を防げる

【契約形態】上記①～④の各段階いずれも、**準委任型の契約**がなじむ

【学習済みモデルの権利帰属および利用条件の設定】ユーザとベンダの利害が対立しやすい点の1つであるが、双方の寄与度等を踏まえつつ具体的に検討・決定することが望まれる

④（行政法）AI・ロボットが個人に関する情報を取り扱う場合、日本の個人情報保護法制との関係で留意すべき点はあるか？

・AI・ロボットにおいて取り扱う個人に関する情報が個人情報保護法制上の個人情報または匿名加工情報（非識別加工情報）であれば、これらの規定に従う必要がある。

11

3. 資料／文献調査（2-3）

⑤（行政法）AI・ロボットがプロファイリングをすることについて注意すべき点はあるか？

・プロファイリングは、現在の日本の法令上は直接規制されていないが、EUの一般データ保護規制（GDPR）等で規制されており、外国の規制動向にも注意する必要がある。

⑥（知的財産法）当社のAIを学習させるために、①特定の第三者が有するデータの集合物を利用する場合、②インターネット上に不特定多数の者がアップロードしたデータを利用する場合、知的財産法上の観点から、それぞれどのような点に気を付ければよいか？

・特定の第三者が有するデータの集合物を利用する場合（①）、不正競争防止法上の「営業秘密」または「限定提供データ」が含まれる可能性があるが、「不正競争」に該当する行為が禁止される。また、当該第三者以外の者の著作物が含まれる可能性があるが、AIを学習させる目的で利用するのであれば、一部の例外を除いて、当該第三者以外の者から個々の許諾を得ずに利用できる。

・インターネット上に不特定多数の者がアップロードしたデータを利用する場合（②）、著作物が含まれる可能性があるが、AIを学習させる目的で利用するのであれば、原則として著作権者の許諾を得ずに利用できる。

⑦（知的財産法）当社のAIに機械学習を行わせて、一定のパラメータ（学習済みモデル）を生成したが、このような学習済みモデルは知的財産法で保護されるか？学習済みモデルの利用にあたっては、どのような点に気を付ければよいか？

・学習済みモデルは、特許法上のプログラムの発明、著作権法上のプログラムの著作物、および不正競争防止法上の営業秘密として保護される場合がある。

・学習済みモデルを利用する際には、他社の学習済みモデルに関する特許権や著作権等を侵害しないことに加え、再利用モデルや蒸留モデルの生成がこれらの権利を侵害しないかについても検討する必要がある。

12

3. 資料／文献調査（3-1）

（3）「機械学習＆ディープラーニングのしくみと技術がこれ１冊でしっかりわかる教科書」山口 達輝／松田 洋之

【概要】

・機械学習＆ディープラーニングの基礎知識、プロセスとコア技術、アルゴリズム、開発の基礎知識などを、例示と図解で分かり易く解説しており、キーワードベースで基礎知識やコア技術をわかりやすく理解できる。

【選択理由】AIシステムのシステム監査をするには、機械学習＆ディープラーニングのアルゴリズムの中で何が行われているか理解している必要があり、機械学習＆ディープラーニングのイメージを直感的に掴めるよう工夫されていることから。

【トピック】 ※以降は全て【出所】「機械学習＆ディープラーニングのしくみと技術がこれ1冊でしっかりわかる教科書」山口 達輝／松田 洋之

①機械学習と特徴量

・機械学習は、ものごとを「分ける」能力の実現のため、「演繹的思考（「AはBである」といった根拠から思考を行う事）」ではなく「統計的思考（「AはBである確率が高い」といった根拠から思考を行う事）」を取り入れて、知能を実現しようとするAIである。

・アルゴリズムの性能を向上させるためには「**どんな特徴量を入れるか**」が重要で、かつその設計は人間自身が考えなければならない、という点がボトルネックとなっている。
⇒ディープラーニングでは、何を特徴量とするべきか**アルゴリズムが自動で抽出**

②機械学習の得意な分野、苦手な分野

（a）過去にデータが存在するか ⇒ 過去のデータがないものは分類も予測もできない
（b）データが十分にあるか ⇒ データ数の少なさが学習のボトルネックとなる
（c）データが定量的か ⇒ **定量的なデータに変換**が難しい課題の解決は苦手
（d）推論の過程がわからなくてもよいか ⇒ **根拠が重要となる**推論が必要な**分野**は苦手

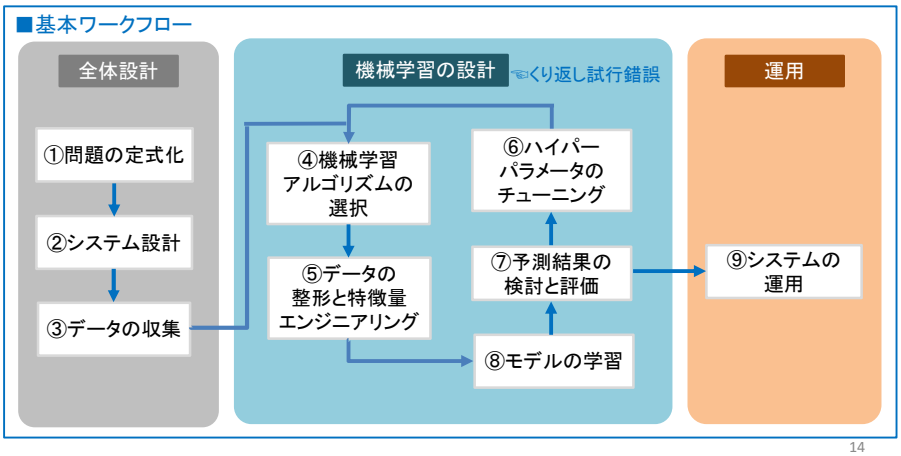
3. 資料／文献調査（3-2）

③機械学習の基本ワークフロー

・機械学習システム開発のポイント

アルゴリズムの選定や機械学習の性能向上のために試行錯誤が多く、プロセス間をまたいだ手戻りが発生しやすいため、**各プロセスに費やす時間を適切に管理する**

⇒**「解決したい問題がそもそも機械学習に向いているのかどうか」を事前に見極めておく**



3. 資料／文献調査（3-3）

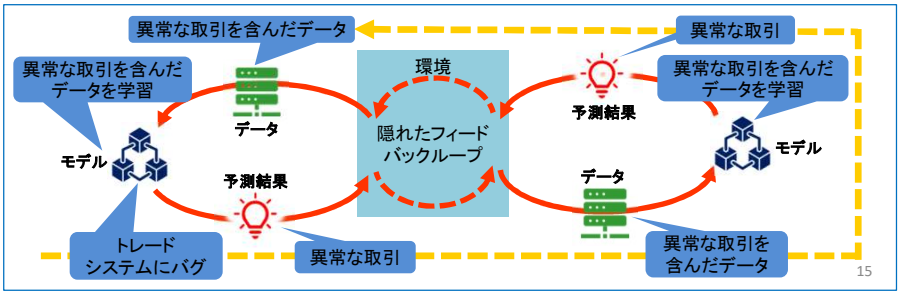
④フィードバックループ

・機械学習のシステムの振る舞いはデータに大きく依存する。そのため、誤りを含んだデータをモデルが学習してしまうと、モデルの出力が意図しないものになる可能性がある。またシステムのうち何かしら一つの要素を変更すると、他のすべての要素も変わってしまうケースがある。
⇒機械学習ではモデルの中身がブラックボックスとなるため、**振る舞いを監視することが重要**

・気を付けたいのが**フィードバックループ**（システムの振る舞いが環境に影響を及ぼし、次に観測するデータが環境から影響を受けて変化してしまう現象）で、システムの振る舞いが徐々に変わっていったり、モデルの更新の頻度が低い場合には、振る舞いの変化に気づくのが遅れる。
⇒システムの使用開始前にその振る舞いを予測するのが難しい

・直接的なフィードバックループの例として予測警備（過去の犯罪のデータをモデルに学習させ、犯罪が多く起こると予測される場所を重点的に警備する警備の方法）がある。

・厄介なのが**隠れたフィードバックループ**で、独立した複数の機械学習システム間で起こる。



15

3. 資料／文献調査（3-４）

⑤ディープラーニングと画像認識

・従来の機械学習では、データをアルゴリズムに入力する前に、データのどこに注目すればよいのかを人間が指定してその値（特徴量）を抽出しなければいけなかった。そのため、画像のような多次元で複雑なデータに対しては、人間が有効な特徴量を設定することが難しかった。
⇒ディープラーニングでは、適切な特徴量を学習の過程で自動的に探し出せるため、人間が気付かないデータのパターンの特徴を利用した処理が可能になった。

・物体検出アルゴリズムにディープラーニングを組み込むことで、より多くの種類の物体をさまざまな状況下で検出できるようになった。物体検出では画像中の物体ごとのラベルを推測しているが、ディープラーニングによる画像処理と⑥の自然言語処理アルゴリズムを組み合わせることで、画像内の複数の物体同士がどのような状況にあるかを説明する文章を自動生成できる（**キャプション生成**）。
⇒物体同士の関係性をコンピュータが認識できるようになり、他分野への応用が爆発的に展開

⑥ディープラーニングと自然言語処理

・**機械翻訳**：単語ごとの翻訳や文法の並べ替え方の設定をやめ、大量の対訳文のデータベースを学習させ、フレーズごとに翻訳させることで、より自然で正確な翻訳を目指した。ディープラーニングを利用した機械翻訳の欠点は、扱える語彙数が小さくなることである。

・**文書要約**：ディープラーニングを使った文書要約タスクで実用的なものの一つにニュースの見出し生成（生成的手法を使った単一文書要約タイプ）がある。文書要約の課題として、㋐機械翻訳よりも多い入力データが必要であること㋑文書要約の正解には絶対的な基準がなく評価が難しいことなどがあげられる。

・**対話システム**：AppleのSiriやGoogleアシスタント、Microsoftの「りんな」のように対話を行うシステムを指す。

・**音声認識**および**音声合成**：Siriなどの最近の対話システムは、テキストのほかに音声による入力も受け付ける。

16

3. 資料／文献調査（3-5）

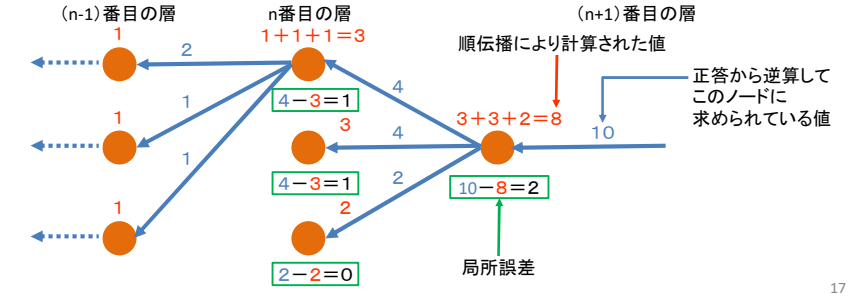
⑦誤差逆伝播法によるニューラルネットワークの学習

・ニューラルネットワークモデルに入力されたデータは、**ノード**と呼ばれる要素の連なりを伝う過程で、各ノードに設定されたパラメータ（重みやバイアス）によりさまざまな処理・変換がなされ、最終層に出力される。この、入力から出力へ情報が伝う流れを**順伝播**と呼び、予測や分類を行う際に利用する。しかし、ニューラルネットワークモデルは、作成した時点ではノードの重みがでたらめに設定されており、出力も正確ではない。

⇒機械学習アルゴリズム同様、学習データによる学習を行う必要があり、代表的手法の１つが誤差逆伝播法である。

・**誤差逆伝播法（バックプロパゲーション）**は、ニューラルネットワークの出力と正答データとの差（誤差）が後ろ（逆）のノードへ伝播するように計算を行い、その重みを調整する手法である。

■誤差逆伝播法

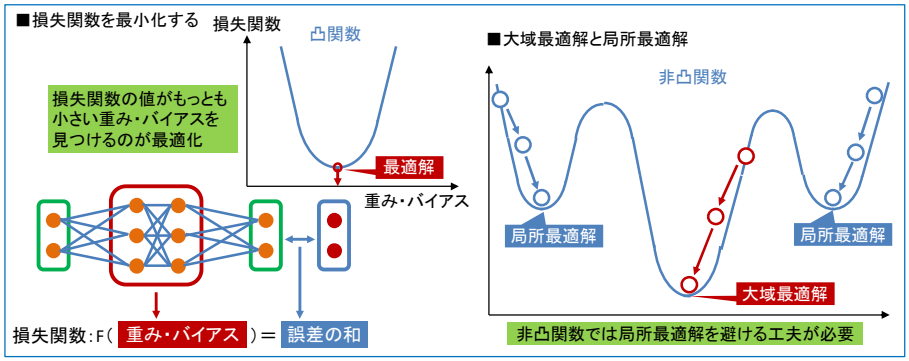


3. 資料／文献調査（3-6）

⑧ニューラルネットワークの最適化

・モデルの最適化は、「**損失関数**（モデルの出力が正答に対してどれだけずれているかを表す値で、ニューラルネットワークの各ノードにおける重みやバイアスから計算できる）」の値が小さくなるようなニューラルネットワークの重みを探し出すことを指す。

・モデルの最適化でもっとも利用されている**勾配降下法**の基本原理（以下参照）



※最新の研究では「ディープラーニングにおいては、ほどよい**局所最適解**（ある範囲の中だけで最もよい解）に収束してもよい性能が得られる」とも言われている。

・最適化において考慮すべき問題には、局所最適解に加えて、**収束の速さ（何回の探索で収束するか）**と**学習率（探索における一歩の大きさ）の設定**がある。一歩が小さすぎるといつまでたっても収束せず、一歩が大きすぎては最適解を通り越してしまうため、慎重に設定する必要がある。⇒これらの問題を解決するために、さまざまな最適化アルゴリズムが開発されている。

3. 資料／文献調査（4-1）

（4）「AIの時代と法」小塚 荘一郎

【概要】

・AIの利用（自動運転、仮想通貨、データ取引、シェアリングサービス、etc.）が普及し、データの価値が増大する時代、取引形態が「モノ（の取引）からサービス（の）取引へ」、取引対象が「財物からデータへ」、取引ルールが「法／契約からコードへ」という変化が生じ、法の世界に大きな変革をもたらす。

・AIの時代に生じるビジネス法の最前線で起きている問題を考え、対処する道筋を描く。

【選択理由】AIを利用していく上でのビジネス利用者と消費者の役割、AIシステムが間違った場合の安全の確保等、監査する上で示唆に富む議論が参考になるとの思いから。

【トピック】 **※以降全て【出所】「AIの時代と法」小塚 荘一郎**

①AI利用の普及・データ価値の増大が経済活動に及ぼす３つの変化

（a）**モノからサービスへ（取引形態）**

（b）**財物からデータへ（取引対象）**

（c）**法／契約からコードへ（取引ルール）**

⇒企業のあり方を変革

②消費者がモノを持たない時代

・CDや書籍が配信サービスになると、その取引は売買契約から、著作物の利用に関するライセンス契約（利用許諾契約）になる。

⇒民法にも著作権法にも規定が置かれておらず、いつ、どのような形での利用が認められるのか、条件が付くのか付かないのかなどは、すべて個別の契約で取り決められる。

・民法の中で、サービスの提供について適用されるのは請負契約と委任契約

⇒請負でも、乗り物を使って人や貨物を移動させる契約は運送契約と呼ばれ、契約の中でどのような取り決めがなされているかにより、サービス内容は大きく変わってくる。

19

3. 資料／文献調査（4-2）

③ AIの間違いと暴走

・想定外の動作をしたり、制御不能に陥ったりして、周囲に損害を発させる危険性

⇒AIに完全に依存するのではなく、セーフティ・ボタンを付けておいて、ぎりぎりの場面では、人間がそれを押して被害の発生を防ぐ。

⇓

・万一の場面で介入を求められる人間は、的確な判断にもとづく対応をとれるのか？

⇒人間の介入が的確なタイミングで行われ、かつ、介入した結果が、AIに任せたままの状態よりも危険を抑える報告に働く仕組みでなければ、安全を確保したシステムとは言えない。

⇓

・医療診断や会計監査の場合、AIが画像データや数値データを読み込むスピードはけた違いに早いので、医師や公認会計士の経験と勘に頼るよりも効率的になるが、AIが病気や粉飾を見落とす間違いをしては困るので、人間が関与してダブルチェックを行う。

⇒AIが詳しく見る必要がないとして捨てたデータを、人間に読めるような画像や文書に変換して、改めて人間がチェックするのでは、データ処理の効率は上がらない。

⇓

・特定のAIシステムが間違いを起こしても、別のシステム（人間の関与も含む）が作動して事故を防止する冗長性のある設計

⇒**人間の判断が介在するといっても、AIによる判断の限界をふまえて、その範囲外にあたる場合の判断フローを適切に設定する。（AIの使われ方を考える責任）**

20

3. 資料／文献調査（4-3）

・職業的にAIを利用する専門家も、AIのプログラムや機械学習などの仕組みについては素人なので、AIシステムの利用者（ユーザー）には、義務や責任などを課すべきではない。
⇒医師は、医療サービスを提供する専門家であり、その専門性にもとづいて適切な医療行為を選択する責任や、医療行為の内容、それを実行した場合のリスクなどについて、患者が理解できるように説明する責任を負っている。その中には、AIをどのように活用し、そのAIが間違った判断をする可能性に対して、どのように対処するかといった「AIの使い方」を考える責任も、含まれる。**AIシステムそのものについて専門的な知識を持っていなくても、**専門業者の活用や、場合によってはAIシステムの専門家を擁する大病院との連携なども含めて、**その適切な利用を組み入れた医療の提供をすることが、専門家としての務め**である。

④プライバシー対「データの活用」

・見える部分の情報を収集・分析しているだけであれば、プライバシーに関し何の問題もないという議論は、おかしい。
⇒見えている情報を収集・分析する中から、見えていない部分、知られたくない情報が浮かび上がってくる。プライバシーに関して問われているのは、適法に収集された情報の使い方について、何も制約はないのかということである。

・消費者Xが、プラットフォームYを利用して、メーカーZの製品を購入したとき、その履歴は、X・Y・Zのうち「誰のもの」なのか？
⇒情報は、特許や著作権によって保護されていない限り、排他的に「誰かのもの」となることはない。商品の購買やプラットフォームの利用などの際の契約（利用規約）にゆだねておくと、利益の配分が不公正になるのではないかという懸念がある。

・GDPRが「データ主体」の権利を宣言しても、一人ひとりに分配される経済的な利益は、意味のある金額にならない。「データ主体」が個人データの利用に同意すれば、ユーザーとして受けられるサービスの条件が有利になるという仕組みが現実的である。 21

3. 資料／文献調査（4-4）

⑤間違わないAIの問題

・差別の禁止には、将来に向けて、社会をよくしていこうという政策が含まれているが、既存のデータに判断を制約された「弱いAI」は、そのような将来に向かっての判断を行う能力を持たない。
⇒AIの判断を最終的に人間が確認し、差別とみられるようなバイアスのかかった判定は修正するというステップを踏むほかはなく、**AIが「間違いなく」行った判定にバイアスがかかっていないかというチェックは、人間の判断がAIとは別の角度から行う。**

⑥デジタル版の「新国際経済秩序」

・EUは、GDPRの制定により、データ（個人情報）の利用に関するルールを厳格にして、利用を拒絶する自由を個人の側に与え、また利用に際して対価が支払われるような仕掛けを導入

・EUは、著作権法ディレクティヴ（指令）を制定（2019年4月）し、大規模なプラットフォームに対して、著作権を侵害するコンテンツを発見し、削除する仕組み（「アップロードフィルター」と呼ばれている）を導入するように義務づけた。

・AIとデータ経済の時代には、「知られたくないものを公開されない」という利益だけでなく、「情報を使って知られたくないことを推測されない」ことの利益が重要になっている。そのような意味のプライバシーは、なおさら、国家に対して守られる必要がある。
⇒プラットフォームには国家と同等の責任があるという議論は、それなりの根拠を持っているが、基本権の保障が確立されていない国に持ち込まれると、企業だけを規制し、国家権力は個人を監視・抑圧するという制度へと容易に転化する危険がある。 22

3. 資料／文献調査（5-1）

（5）「AI・データの利用に関する契約ガイドライン1.1版」経済産業省

【概要】

・AI技術を利用したソフトウェア（特定目的のための特化型AI技術を利用したソフトウェア）について、その特性を踏まえた上で、開発・利用契約を作成するにあたっての考慮要素、トラブルを予防する方法等についての基本的な考え方を解説するとともに、具体化したモデル契約を提示している。

【選択理由】（３）で取り上げた「AI・ロボットの法律実務Q&A」の中で、「AI・データの利用に関する契約ガイドライン」をベースに解説している部分があり、AI利用の契約面からの課題に対する理解を深め、整理したいという思いから。

【トピック】 **※以降は全て【出所】「AI・データの利用に関する契約ガイドライン1.1版」 経済産業省**

①AI技術を利用したソフトウェアの開発・利用をめぐる契約の現状

・AI技術を利用したソフトウェアを用いた事業に関する類型は、（a）ユーザがベンダに依頼して、学習済みモデルの生成を行う「開発」型と、（b）AI技術を利用したサービスを提供するサービス利用」型が想定される。学習済みモデルの生成では、ユーザがベンダに対してデータを提供し、ベンダがそれを学習用プログラムに学習させることが多い（ただし、ベンダがデータを提供することもある。）。AI技術を利用したサービスでは、ユーザがベンダに対してデータを提供することが多い。学習済みモデルの生成者であるベンダと、生データまたは入力データの提供者であるユーザとの間において、知的財産権の帰属や利用条件等について、利害が対立することがしばしばある。また、学習済みモデルまたは AI技術を利用したサービ スの品質に関して、当事者の利害が対立することがある。その結果、事業上の優越関係や技術的な知識の格差等を背景として、「オール・オア・ナッシング」の一方的な契約条項が押しつけられることもある。

⇒一見、当事者の利害が対立するように見えても、学習済みモデルの特性と法律上のルールの内容を理解することで、合理的な条項に合意することができる場合もある。 23

3. 資料／文献調査（5-2）

②契約の検討に向けた視点

・学習済みモデルの生成・利用を目的とする契約の具体的な内容を検討するにあたっては、**AI技術の特性を前提とした上で、事業を進める際に、各当事者が何を守る必要があるか、またリスク要因がどこにあるか、契約による合意の対象を確定することが重要**である。

（注）AI技術の特性

（a）学習済みモデルの内容・性能等が契約締結時に不明瞭な場合が多い

（ｱ） 事前の性能保証が性質上困難である

（ｲ） 事後的な検証等が困難である

（ｳ） 探索的なアプローチが望ましい

（b）学習済みモデルの内容・性能等が学習用データセットによって左右される

（c）ノウハウの重要性が特に高い

（d）生成物に更なる再利用の需要が存在する

・各当事者の立場や考え方の違い

ユーザ側
- □ 開発費を支払い、学習済みモデル生成のための学習に用いるために価値あるデータ・ノウハウを提供したのだから、学習済みモデルに関する権利は全部自社のものとしたい。
- □ 学習済みモデルを競合事業者に使われたくない。
- □ 自社のデータ・ノウハウを外部に流出させたくない。
- □ 学習済みモデルやこれを用いたシステムは一定レベルのものを完成・納品してもらいたい。
- □ 自らのデータを使って追加学習させて学習済みモ デルの精度をさらに上げたい。

ベンダ側
- □ 自社の研究・開発に関する事業自由度を確保したい。
- □ プログラムやシステムに関する権利は、開発主体である自社に帰属してしかるべきである。
- □ 学習済みモデルを横展開して一定の範囲で他社にも提供したい。
- □ 追加学習して精度を上げた学習済みモデルを生成したい。
- □ そもそもユーザの求める目的に合致する学習済みモデルを作成できるかどうかはやってみないとわからない。
- □ 学習済みモデルの完成や未知の入力（データ）に対して性能の保証はできない。

・当事者間で問題が生じうる事項

学習済みモデルの生成・利用に関する契約を締結するに際しては、**権利帰属・利用条件の設定と、責任の所在を明確化することが重要**である。 24

3. 資料／文献調査（5-3）

③学習済みモデル生成の場合の契約の法的性質

	アセスメント	PoC	開発	追加学習
目的	一定量のデータを用いて学習済みモデルの生成可能性を検証する	学習用データセットを用いてユーザが希望する精度の学習済みモデルが生成できるかを検証する	学習済みモデルを生成する	ベンダが納品した学習済みモデルについて、追加の学習用データセットを使って学習をする
成果物	レポート等	レポート／学習済みモデル（パイロット版）等	学習済みモデル等	再利用モデル等
契約	秘密保持契約書等	導入検証契約書等	ソフトウェア開発契約書	多様なものが想定され、保守運用契約の中に規定することや学習支援契約または別途新たなソフトウェア開発契約を締結することが考えられる

・アセスメント段階：重要なことは、事業上の課題およびKPIの設定で、いずれも事業内容に依存することから、ユーザの責任において実施され、ベンダはそれを支援する役割分担が実情に即している。学習済みモデルの生成に際しては、ユーザの積極的な関与が必要不可欠である。
・PoC段階：PoC段階では、様々な業務が対象となるため、PoC段階の契約については、その対象範囲や対象期間を合意しておくことが重要である。PoC段階は、学習済みモデルの生成が試行錯誤を不可避的に伴うことから、1回で完結せず、複数回実施されることも少なくない。PoC段階では、その後の開発段階への移行が想定されているため、それぞれの段階で統一的に取り扱うべき事項があるか、あるいは、各段階で個別に取り扱うべき事項があるかを整理しておく。
・開発段階：前記②（a）～（d）学習済みモデルの特性から、契約締結時までに仕様や検収基準を確定することは難しく、未知の入力（データ）に対して、学習済みモデルがユーザ・ベンダのいずれもが想定しない挙動をしないことの保証をすることも困難であるため、準委任型の契約がなじむ。
・追加学習段階：学習済みモデルを生成したベンダが追加学習支援をすることもあれば、全く別のベンダが実施する場合もある。保守運用とセットでなされることも考えられる。

⇒**学習済みモデル生成の各段階には、**具体的な学習済みモデルの完成を約束する請負型の契約ではなく、**一定の検証や開発といった役務の提供を目的とする準委任型の契約がその実態になじみやすい。**

25

4. 課題（1/8）

（1）AIとは何なのか（定義）？

①人工知能（AI）のイメージ（日米）

【出所】平成28年版 情報通信白書【総務省資料】

調査への回答者（回答者数）	複数回答における回答割合が1番と2番となった選択肢（回答割合）
日本の就労者（1,106人）	コンピュータが人間のように見たり、聞いたり、話したりする技術（35.6%）
	コンピュータに自我（感情）を持たせる技術（27.4%）
アメリカの就労者（1,105人）	人間の脳の認知・判断などの機能を、人間の脳の仕組みとは異なる仕組みで実現する技術（42.3%）
	コンピュータが人間のように見たり、聞いたり、話したりする技術（36.9%）

⇒日米の就労者の抱くAIのイメージは、日米双方で、人間の知覚や会話の代替イメージが多いが、加えてアメリカでは、人間の脳の認知・判断の代替イメージが勝っている。

26

4. 課題（2/8）

②国内の主な研究者による人工知能（AI）の定義　【出所】平成28年版 情報通信白書［総務省資料］

研究者	所属	定義
中島秀之	公立はこだて未来大学	人工的につくられた、知能を持つ実態。あるいはそれをつくろうとすることによって知能自体を研究する分野である
武田英明	国立情報学研究所	
西田豊明	京都大学	「知能を持つメカ」ないしは「心を持つメカ」である
溝口理一郎	北陸先端科学技術大学院	人工的につくった知的な振る舞いをするためのもの（システム）である
長尾真	京都大学	人間の頭脳活動を極限までシミュレートするシステムである
堀浩一	東京大学	人工的に作る新しい知能の世界である
浅田稔	大阪大学	知能の定義が明確でないので、人工知能を明確に定義できない
松原仁	公立はこだて未来大学	究極には人間と区別が付かない人工的な知能のこと
池上高志	東京大学	自然にわれわれがペットや人に接触するような、情動と冗談に満ちた相互作用を、物理法則に関係無く、あるいは逆らって、人工的につくり出せるシステム
山口高平	慶應義塾大学	人の知的な振る舞いを模倣・支援・超越するための構成的システム
栗原聡	電気通信大学	人工的につくられる知能であるが、その知能のレベルは人を超えているものを創造している
山川宏	ドワンゴ人工知能研究所	計算機知能のうちで、人間が直接・間接に設計する場合を人口知能と呼んで良いのではないかと思う
松尾豊	東京大学	人工的につくられた人間のような知能、ないしはそれをつくる技術。人間のように知的であるとは、「気づくことのできる」コンピュータ、つまり、データの中から特徴量を生成し現象をモデル化することのできるコンピュータという意味である

27

⇒ **AIには、確立した学術的な定義や合意がありません。**

4. 課題（3/8）

③本プロジェクトでのAIの定義

・人工知能学会では「知的な機械、特に、知的なコンピュータプログラムを作る科学と技術」と説明されているが、研究者によって定義は異なっている。

・『知性』や『知能』自体の定義がないから、人工的な知能を定義することも困難。

⇒取りあえず**「人間と同じような知的処理を行うことのできる技術や機械」**と定義。

また、一般的なシステムにAIの要素を組み込んだシステムをAIシステムと呼ぶ。

（２）AIの大分類　（「総務省ICTスキル総合習得教材」を基に加筆）

分類Ⅰ	説明	イメージ・事例	分類Ⅱ	説明
特化型AI	限定された領域の課題に特化して自動的に学習、処理を行うAI	**「人間が知能を使ってすることを機械にさせようとする立場」（実際の研究のほとんどはこちら）** 画像認識や音声認識、自然言語処理や医療診断などの技術を持つAI ビジネス領域で広く活用されている	**弱いAI**	人間の知性の一部分のみを代替し、特定のタスクだけを処理するAI
汎用型AI	特定の課題にのみ対応するのではなく、人間と同じようにさまざまな課題を処理可能なAI	**「人間の知能そのものをもつ機械を作ろうとする立場」** 人間のような問題処理能力（想定外の出来事が起きた場合でも、これまでの経験に基づいて総合的に判断し、問題を解決出来る）を持つAI 最終的な将来像を表現した言葉	**強いAI**	人間のような自意識を備え、全認知能力を必要とする作業も可能なAI

28

4．課題（4/8）

（3）AIの利活用

①AI利活用を想定した産業

（「総務省ICTスキル総合習得教材」を基に加筆修正）

大分類	業種	業務	要求精度
農業、林業	農業	耕うん・整地、播種・育種、追肥・除草、収穫、調製、見張り	中
建設業	建設	建設現場での各種作業	中
	住宅ﾘﾌｫｰﾑ	解体、搬入、塗装、設置、ﾏｰｹﾃｨﾝｸﾞ	
製造業	自動車・自動車部品	生産	中
		自動運転、事故防止	高
	医薬品	試行錯誤による製薬	高
電気・ガス・熱供給・水道業	電力	点検、建設	中
		異常監視、廃炉作業	高
情報通信業	通信業	設備保守、使用状況のﾓﾆﾀﾘﾝｸﾞ	中
		異常監視	高
	BtoC-EC	ﾈｯﾄ販売の広告・推薦	中
	通販	ﾈｯﾄ販売	中
運輸業、郵便業	鉄道	設備保守、移動からの広告表示	中
		異常監視、事故防止	高
	物流	積み替え、運転、戸口配送	中

大分類	業種	業務	要求精度
卸売業、小売業	スーパー・百貨店	陳列、補充、会計、清掃、防犯、顧客行動分析、ﾏｰｹﾃｨﾝｸﾞ	中
	ｺﾝﾋﾞﾆ・ﾄﾞﾗｯｸﾞｽﾄｱ	陳列、補充、発注、防犯、顧客行動分析、販売促進	
	アパレル	陳列、補充、顧客行動分析、ﾏｰｹﾃｨﾝｸﾞ	
	家電小売	陳列、補充、在庫管理	
金融業、保険業	銀行	ﾃﾞｰﾀ解析、ﾈｯﾄ銀行、ｺｰﾙｾﾝﾀｰ	中
	生命保険	顧客毎の料率計算、健康管理	
	損害保険	ﾈｯﾄ保険	
不動産業、物品賃貸業	不動産	防犯・監視付加価値、物件検索	中
学術研究、専門・技術サービス業	広告	視聴者の反応分析、ﾈｯﾄ広告	中
宿泊業、飲食サービス業	外食	調理、接客、ﾏｰｹﾃｨﾝｸﾞ	中
	中食	食品加工、配送	
生活関連サービス業、娯楽業	旅行	ﾈｯﾄ販売	中
医療、福祉	医療	画像診断、VR、見守り	高
		健康アプリ	中
	介護	見守り、移動、ﾄｲﾚの世話、ｺﾐｭﾆｹｰｼｮﾝｱﾌﾟﾘ	中

29

4．課題（5/8）

②AIの利活用が望ましい分野（有識者27人アンケート）【出所】平成28年版 情報通信白書［総務省資料］

AIの利活用が望ましい分野	割合(%)
生体情報や生活習慣、病歴、遺伝等と連動した、健康状態や病気発症の予兆の高度な診断	81.5
路線バスやタクシー等の高度な自動運転	81.5
渋滞情報や患者受入可能な診療科情報等と連動した、緊急車両の最適搬送ルートの高度な設定	77.8
道路や鉄道などの混雑状況等と連動した、交通手段間での高度な利用者融通や増発対応	74.1
監視カメラ映像や不審者目撃情報等と連動した、犯罪発生の予兆の高度な分析	70.4
高度かつリアルタイムの需要予測や製造管理等によるサプライチェーンの最適化	66.7
未知のサイバー攻撃や内部犯行等による不正アクセスや、不正送金などの金融犯罪の高度な検知	66.7
高度な意味理解や感情認識等によるコンピュータと人間の対話の高度化	48.1
利用者の嗜好やメール履歴、発信元等と連動した、迷惑メールの高度かつ自動的な削除	44.4
市場の値動き等と連動した、金融資産の高度かつ自動的な運用による利回りの最大化	37.0
信用供与先の財務状況等と連動した、最適な融資額の算定による貸倒れ損失の回避	37.0
優良顧客の優遇や感動体験の付与、需給に見合う価格設定等による、顧客の囲い込みや満足度向上	25.9
その他	37.0

・健診の高度化や公共交通の自動運転、救急搬送ルートの選定、交通混雑・渋滞の緩和など、社会的課題の解決が期待される分野で、AIの利活用ニーズが相対的に高い。
・金融やマーケティング、コミュニケーションといった産業や個人の生活に関わる分野では、AIの利活用ニーズが相対的に低い。

30

4．課題（6/8）

（3）AIを利用した開発の必要性はどうやって判断するのか？

◆ユーザがAIをよくわかっていない場合がある

①日本企業のAIに関する10の誤解　【出所】日本におけるテクノロジのハイプ・サイクル：2016年［ガートナージャパン資料］

1	**すごく賢いAIが既に存在する。**
2	**IBM Watson のようなものや機械学習、深層学習を導入すれば、誰でもすぐに「すごいこと」ができる。**
3	AIと呼ばれる単一のテクノロジーが存在する。
4	AIを導入するとすぐに効果が出る。
5	「教師なし学習」は教えなくてよいため、こちらの方が良い。
6	ディープラーニングが最強である。
7	アルゴリズムをコンピュータ言語のように選べる。
8	誰でもがすぐに使えるAIがある。
9	AIとはソフトウェア技術である。
10	結局、AIは使い物にならないため意味がない。

多くの誤解が1.と2.に見られる。

1.について、「経営者やテクノロジーにそれほど詳しくない人は、AIによってさも"今、人間と同様のことができる""今すぐにすごいことができる"と捉えてしまう傾向がある。AIの研究者は、現時点で"人間と同様の知能"を実現できているテクノロジーは存在しないことを当たり前のこととして認識している。AIに関しては、遠い将来の話と、現在の話、数年後の話といったことを明確に分けて捉えるべき」と提言。

2.について、「企業は、AIのようなものを導入すれば誰でもすぐにすごいことができるというのは誤りであることを、まずは理解する必要がある。"AIのようなもの"とは、人間に例えれば赤ちゃんか子供であると捉えておくべきで、うまく育てるためにも、育てる人のスキルが求められる」と指摘。

31

4．課題（7/8）

②AIの特性を理解して使いこなす

①の誤解を払底するには、実証実験（PoC）などスモールスタートでAIへの理解を深め、実際にAIを使用してみることで、その可能性や課題を明らかにすることが必要である。

⇒AIの特性を理解した上で、人間が使いこなす必要がある。

（A）AIにはできないことがある

・AIは自発的に目的や課題を設定できない。AIを使って何を実現したいのか、目標や課題を設定し、それに応じた入出力を定義するのは人間の仕事である。

・AIに十分に学習させるには、大量の偏りのないデータを用意する必要がある。

・AIが出した判断が人間の感覚と乖離している場合、AIが出した判断を評価、解釈し、その結果を踏まえて行動し、場合により他の人を巻き込むことも人間の役割として残る。

（B）データ学習なくしては使えない

・学習によって得た知識を一般化することで未知の局面にも対応できることがAIの大きな強みである。学習フェーズでは、目的や課題の解決となるアウトプットを得るために仮説を設定し、どんなデータを入力すればよいか、何のアルゴリズムを用いるのがよいか試行錯誤しながら検証を繰り返して学習済みモデルを導き出し、実際の本番データに適用することで、期待するアウトプットを得る。精度を高めるためには、現場やユーザの生の声から学習データを作成する必要がある。

・学習サイクルを運用フローに組み込むに際して、予期せぬ方向に成長する恐れがあるので、学習フェーズと本番稼働は分けておき、何かしらの静止点をもって本番に適用するのが安全である。

（C）AIは間違えることもある

・AIは、未知の本番データに対して、どのような答えを出すかは事前に予知できないため、テストによる動作保証に限界があり、AIが間違った答えを出すこともあり得る。そうした事態に備えて人間が運用面でカバーできる体制を作っておく必要がある。

32

4．課題（8/8）

（4）AIを利用した開発のゴールはどうやって決めるのか？

◆AIでは運用で、開発でゴールは無いので、ある目標で止める必要がある

・オーナーは、AIにどの程度の精度を要求するか明らかにしておく。業務内容および運用体制まで踏み込んで、コストや時間と精度のバランスを取ることを考慮する。

◆オーナーはAIの成果物を生成することを止める権利を行使すべき

・開発を複数段階に分け、予定した性能を発揮できないことが判明した段階で開発を中止することにより、それ以上の損失拡大を防ぎ、リスクヘッジを図る。

（5）AIを導入するリスクはどれくらいあるのか？

◆人間は結果を見て正しいと判断出来るのか？

⇒AIが出した判断が人間の感覚と乖離している場合、判断の根拠が不明瞭な中で、最終的にどう振る舞うか責任をもって判断できないリスクがある。

◆以下「ディープラーニングの特徴」は、AIシステムのリスクマネジメント、ガバナンス上の問題点を明示している。

【ディープラーニングの特徴】 【出所】AI時代における監査の取り組みとポイント：2019年10月度【ISACA月例会資料】

（A）なぜ結果が導き出されたかについて知ることができない（説明ができない）
（B）人間が期待する結果とは異なる結果を導き出す恐れがある ⇒ 学習データによる
（C）バグゼロでも人間の期待に沿わない結果が導き出される可能性がある
⇒ 学習データによる
（D）過学習発生により、あらたなデータに適応できない場合がある
（E）稼働後の品質保証が困難 ⇒ **常にモニタリングしておく必要がある**

◆ディープラーニングは、結果が導出された経緯を説明出来ない

⇒結果に対する判断・責任が曖昧となり、コンプライアンス、ガバナンスの視点から、当事者責任を果たせないリスクがある。

33

5．AIと監査

（1）AIを利用する分野によって監査手法は変わってくるのでは？

◆対象とするAIシステムを特定しないと、システム監査の対象にならない

・要求精度の高い業種とそれほどでもない業種では、リスクの程度も異なる。

◆危険を扱った成果物

（2）AI監査でも専門監査人が必要とされるのではないか？

◆AIとはどういうものかを知った上で監査するべきではないか？

・監査人は、監査サービスを提供する専門家であり、その専門性にもとづいて適切な監査方法を選択する責任や、監査基準、監査手続き、AIシステムを開発・運用した場合のリスクなどについて、システムオーナーが理解できるように説明する責任を負っている。その中には、AIをどのように活用し、そのAIが間違った判断をする可能性に対して、どのように対処するかといった「AIの使い方」を考える責任も、含まれる。**AIシステムそのものについて専門的な知識を持っていなくても、AIシステムの専門家の活用なども含めて、AIシステムを総合的かつ客観的に点検・評価し、関係者に助言・勧告することが、専門家としての務めである。**

◆何をもってどういう条件で専門監査人と言えるのか？

（3）監査人の倫理、善管義務とは？

◆監査人自身が責任を負えないのであれば、監査人は辞任するべきである

◆最終判断は人間がやるべき

34

６．今後の展開

・これまでに分かったことは、
「AIの特性を理解し、出来ることと出来ないことを見極めた上で、オーナーとともにAIシステムのリスクを適切にコントロール・運用していく必要がある」のではないかということです。
それには助言型監査の方がなじむのではないかということで、そのためには、ここは必要で、ここは不要というものをあげていく必要があると考えます。
・AIの監査基準みたいなものが必要ではないかと考え、目次だけでも作れないかと思いましたが、そこまでやるケースが少ない今の段階の状況では、作成するのは難しい。
・今後は「ＡＩシステム」の監査事例について調査・検討を行っていき、AIシステムを監査する上でのポイントをまとめて行きたいと思っています。

35

ご清聴、ありがとうございました。

36