報告書 2020

～「安心・安全で信頼性のある AI の社会実装」に向けて～

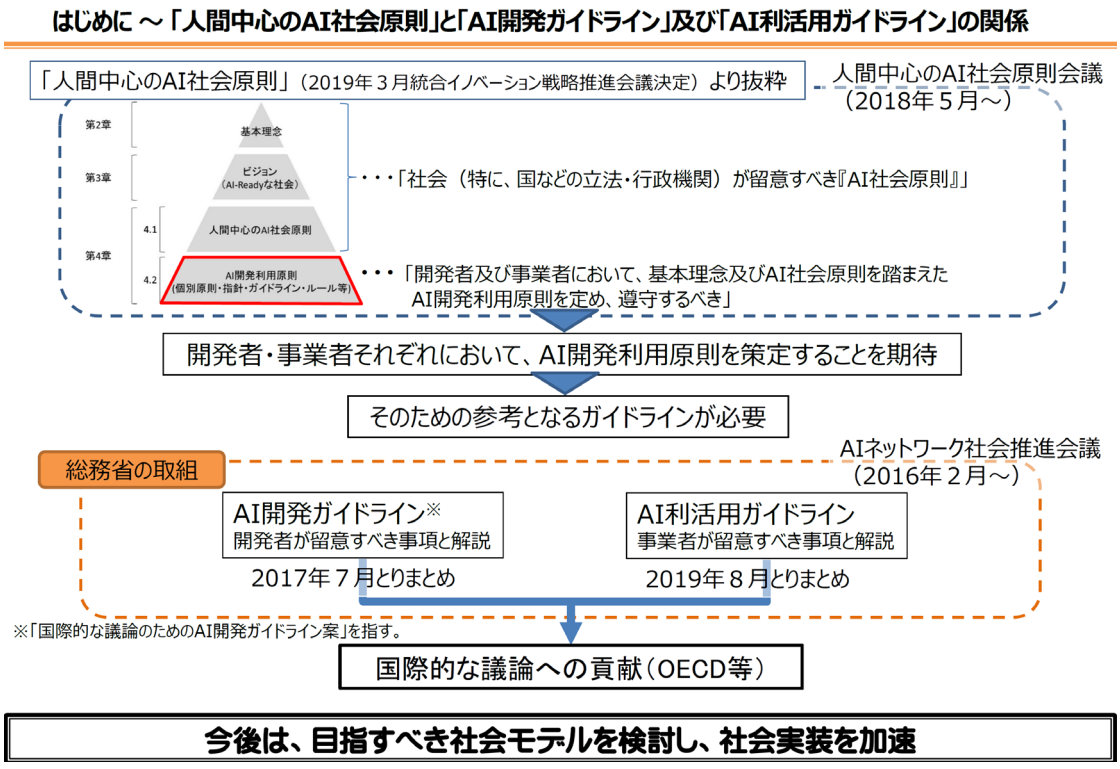
令和２年７月２１日

AI ネットワーク社会推進会議

目次

はじめに

AI ネットワーク社会推進会議（以下「本推進会議」という。）では、2019 年（令和元年）8 月に「AI 利活用ガイドライン」を含む『報告書 2019』をとりまとめ公表した。「AI 利活用ガイドライン」は、AI の利用者が留意することが期待される事項を整理したものであり、2017 年（平成 29 年）7 月に本推進会議において AI の開発者が留意することが期待される事項を整理した「AI 開発ガイドライン」[1]と対をなすものである。そして、「人間中心の AI 社会原則」（2019 年（平成 31 年）3 月 統合イノベーション戦略推進会議決定）において、開発者及び事業者が、同原則で定める基本理念及び AI 社会原則を踏まえた AI 開発利用原則を定め、遵守することが期待されているなかで、両ガイドラインは、開発者及び事業者が自らの原則等を策定する際に参考となるものと位置づけられるものである [2]。



また、そもそも開発者及び事業者が自らの原則等を策定する趣旨は、AI の利活用が期待される一方で、人々の AI に対する不安などが、AI の開発及び利活用の促進や AI ネットワ

[1] 本推進会議が 2017 年（平成 29 年）7 月にとりまとめた『報告書 2017』において、AI の開発者が留意することが期待される事項を整理した「国際的な議論のための AI 開発ガイドライン案」を指す。

[2] 『報告書 2019』「はじめに」を参照。

ーク化の健全な進展の阻害要因になるのではないかと懸念されていることを踏まえ、人々のAIに対する不安を取り除き、信頼を醸成する取組を進めることにある。そこで、こうした開発者及び事業者が自らの原則等を策定する本来の趣旨を踏まえると、両ガイドラインは、「安心・安全で信頼性のある AI の社会実装」を積極的に推進していくための参考となるツールとしての役割を担っていくものと考えられる。

こうした認識のもと、本推進会議事務局において両ガイドラインの周知に努めるとともに[3]、『報告書2019』の「**第２章　ＡＩ利活用ガイドライン策定の考え方　４．今後の展開**」及び「**第３章　今後の課題**」も踏まえ、原則等の策定も含め様々なステークホルダと「安心・安全で信頼性のあるAIの社会実装」について意見交換を行ってきた。そして、こうした意見交換も踏まえ、「安心・安全で信頼性のある AI の社会実装」に向けて必要となる論点を提示し、AIに関して意欲的な取組を行っている方々から本推進会議議長によるヒアリング（以下「本ヒアリング」という。）という形式によりヒアリング及び自由闊達な議論を重ねてきた。その主な論点は以下のとおりである。

（１）開発者及びAIサービスプロバイダー（以下「開発者等」という。）について

ア　開発者等については、自らのAI原則等[4]の策定がそれぞれ特長を有する形で進められる中で、「安心・安全」、「信頼性」ということが重視されていると考えられる。そこで、さらに開発者等におけるAI原則等の策定を促進する観点から、あらためてAI原則等の策定の意義やそれを安心・安全で信頼性のあるAIの開発・利活用にどのように活かすことができるか。

イ　AI原則等の運用も含め、安心・安全で信頼性のあるAIの開発等に必要なガバナンス体制（自己点検・自己評価の仕組み／外部による評価の仕組み）にはどのようなものが考えられるか。

（２）ビジネス利用者について

ビジネス利用者には様々な類型があるが、AI の利活用を進めるにあたり課題となっていることは何か。また、その課題解決に必要な取組は何か。

（３）消費者的利用者について

消費者的利用者については、AI 利活用ガイドラインにおいて、参考という位置づけで記述されているが、「安心・安全で信頼性のある AI の社会実装」を積極的に推進していくと

---

[3] 周知活動の一環として、総務省広報誌2019年（令和元年）12月号において、「AI、どう使う？知って安心　AI利活用」を特集として掲載している。次に掲げるURLのウェブサイトに所掲。
<https://www.soumu.go.jp/main_content/000656829.pdf>

[4] 事業者によって、AI原則、AIコミットメント、AI指針等名称は様々であるが、総称として「AI原則等」と呼ぶこととする。

いう観点からは、消費者的利用者が安心してAIを利活用し、その便益を享受できる取組が必要となってくる。そこで、

ア　消費者的利用者に関する取組をどのように進めていくことが考えられるか。

イ　とりわけ、高齢者・障害者がAIを利活用することにより、加齢あるいは障害を有することに伴う不便を解消することで、誰もが等しく自己実現を図れるようにすることは、人間中心のAI社会を実現する上で最も重要な取組の一つであると考えられる。そこで、高齢者・障害者にとっての安心・安全で信頼性のあるAIの社会実装の取組としてどのようなものが考えられるか。

（4）安心・安全で信頼性のあるAIの社会実装のための環境整備

関係する主体に関する取組の観点とは別に、環境整備として、

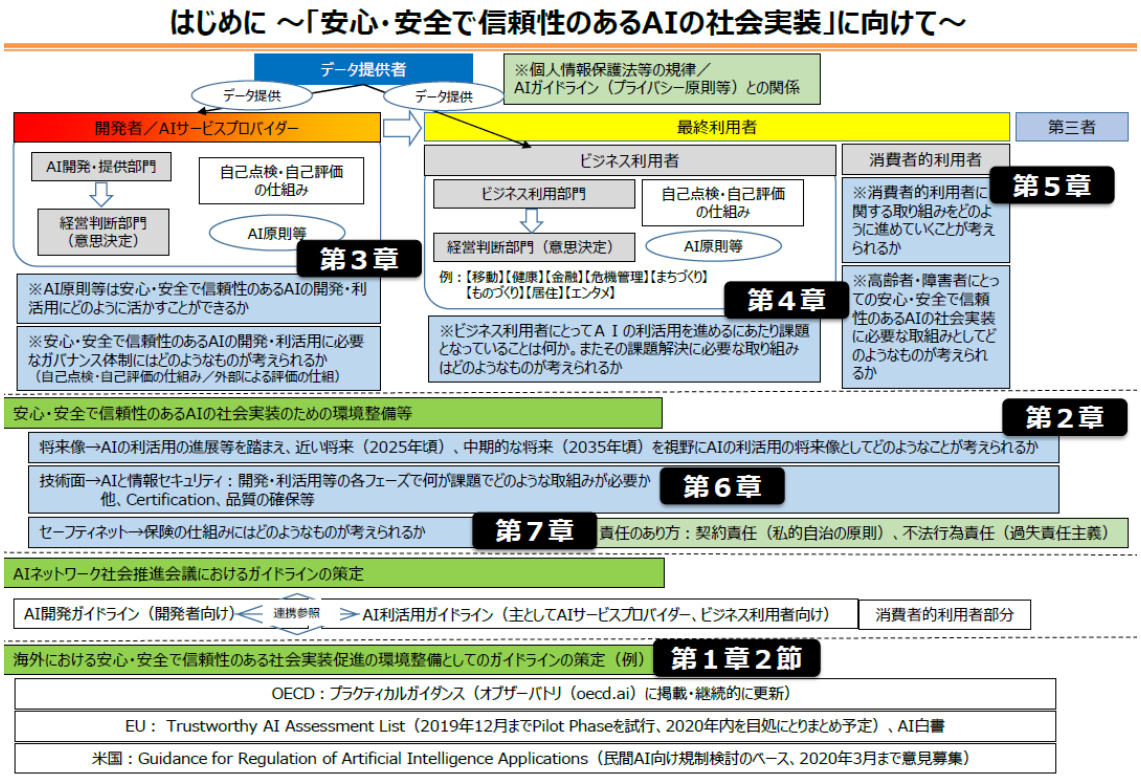
ア　技術的な観点から「AIと情報セキュリティ」について、開発・利活用等の各フェーズで何が課題でどのような取組が必要か。

イ　セーフティネットの観点から、保険の仕組みとしてどのようなものが考えられるか。

以上の論点を中心として行ってきた本ヒアリングについて、ヒアリング対象者の許諾のもと、その範囲内でヒアリング内容及び議論について整理し、必要なとりまとめを行っている。

また、「安心・安全で信頼性のあるAIの社会実装」を進めるためには、AIの社会実装が進展した先にある社会像のシナリオについて分析し提示することで具体的なイメージを共有することが有益である。そこで、将来像として、AIの利活用の進展等を踏まえ、近い将来（2025年頃）、中期的な将来（2035年頃）を視野にAIの利活用の将来像を分析し提示することで具体的なイメージの共有を試みることとした。

さらに、「安心・安全で信頼性のあるAIの社会実装」の取組の検討にあたっては、AIネットワーク化の性質から国際的な連携が不可欠であることから、経済協力開発機構（OECD）、欧州連合（EU）、米国等における取組を把握するとともに必要に応じた対応を行っていくことが重要である。こうした観点から、本ヒアリングにおいて海外及び国際的な議論の動向についてもヒアリングを行い、「安心・安全で信頼性のあるAIの社会実装」の取組の検討に必要な情報についてとりまとめている。



本とりまとめの考え方は以上のとおりとなっている。なお、本とりまとめは個別具体的かつ意欲的な取組等のヒアリングを踏まえ行っているものであり、「安心・安全で信頼性のある AI の社会実装」にとって必要な取組を必ずしも網羅するものとはなっていないとも考えられる。しかしながら、本とりまとめを通じて、こうした個別具体的かつ意欲的な取組等を広く紹介するとともに、「安心・安全で信頼性のある AI の社会実装」の参考となる取組として広く共有されることは、AI の開発・利活用を真摯に検討している関係者にとっては大変有益な情報になると考えている。本とりまとめが広く参考とされ、関係者において必要な取組が進められることで、我が国における「安心・安全で信頼性のある AI の社会実装」の推進に貢献していくことを願っている。

第１章 AI ネットワーク化をめぐる最近の動向

本章においては、主として『報告書 2019』公表後の AI ネットワーク化をめぐる動向（特に、「安心・安全で信頼性のある AI の社会実装」に関連する動向）を概観する[5]。

１． AI と COVID-19[6]対策

今般、COVID-19 の拡大等に対処すべく様々な取組が行われているが、その中には AI に関連するものも多く存在する。例えば、EU の AI Alliance では”Join the AI-ROBOTICS vs COVID-19 initiative”[7]と題し、関連する取組等を紹介している。また、OECD では、COVID-19 に対する AI 技術の活用について、ベースとなる技術及び応用分野について分類[8]したものを示している。本節では、同分類を参考に、各国・地域で行われているそれぞれの取組[9]

---

[5] 『報告書 2019』公表以前の AI ネットワーク化をめぐる動向については、『報告書 2019』第 1 章参照。

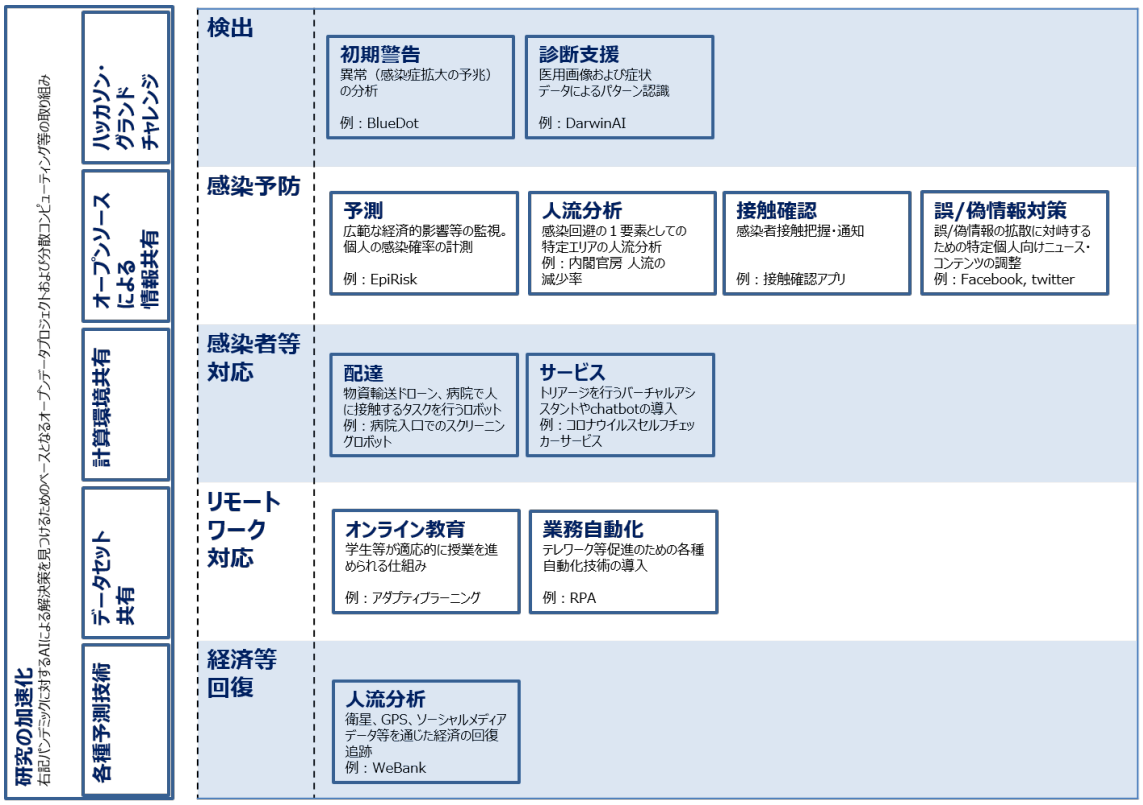
[6] (2019 年)新型コロナウイルス感染症のこと。”COVID-19”は 2019 年（令和元年）に発生した Coronavirus disease の略称として、2020 年（令和 2 年）2 月、世界保健機関（WHO）が命名。
<https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf>

[7] Join the AI-ROBOTICS vs COVID-19 initiative of the European AI Alliance
<https://ec.europa.eu/digital-single-market/en/news/join-ai-robotics-vs-covid-19-initiative-european-ai-alliance>

[8] 当該分類は OECD の以下文書に含まれるものに対する仮訳である。なお、この仮訳は OECD によって作成されたものではなく、OECD の公式な翻訳とは見なされない。また、OECD は、本翻訳の内容または誤りに対して責任を負わない。
OECD/Using artificial intelligence to help combat COVID-19,
<https://read.oecd-ilibrary.org/view/?ref=130_130771-3jtyra9uoh>
“This translation was not created by the OECD and should not be considered an official OECD translation. The OECD shall not be liable for any content or error in this translation.”

[9] 上記以外、例えば以下を参照。
Gartner: Top five areas CIOs can use AI to combat COVID-19
<https://remoteworkertech.asia/story/gartner-top-five-areas-cios-can-use-ai-to-combat-covid-19>
Ledge.ai：新型コロナと戦う AI 技術－感染症患者検知や行動分析など 23 事例を紹介
<https://ledge.ai/aicompany-fighting-covid19/>
AInow：【事例を元に解説！】AI は新型コロナ感染拡大にどう向き合うか
<https://ainow.ai/2020/05/14/222586/>
新型コロナウイルス感染症対策 テックチーム Anti-Covid-19 Tech Team（第 2 回）
テックチーム 現在進行中のプロジェクト一覧
<https://cio.go.jp/sites/default/files/uploads/documents/techteam_20200421_01.pdf>
新型コロナウイルス感染症対策 テックチーム
<https://cio.go.jp/techteam>
新型コロナウイルス感染症（COVID-19）に関するオープンコラボレーション

を下図のように分類した上で概観する[10]。



（１） 研究の加速化

後述の（２）から（６）に述べる AI を活用した COVID-19 への対処等に際し、考え得る研究加速化の取組及び手段について以下に述べる。なお、我が国の研究機関を中心とした大学・公的機関での取組については、人工知能研究開発ネットワークのホームページ[11]に詳しくまとめられているので参照されたい。

ア. （構造・創薬などの）各種予測技術

ウイルスに関連する構造の同定及び創薬の加速化のために AI が使われている。例えば、Deepmind 社は、自社の最新 Alphafold システムを利用して COVID-19 を引き起こすウイルスに関連するタンパク質構造を予測したことを発表している[12]。

<https://github.blog/jp/2020-04-02-open-collaboration-on-covid-19/>

10 なお、ここに示す取組は分類を含め一例であり、すべての取組を網羅的に示したものではない。

11 人工知能研究開発ネットワーク「新型コロナウイルス感染症対策に係る AI を活用した取組」
<https://www.ai-japan.go.jp/COVID19>

12 DeepMind: Computational predictions of protein structures associated with COVID-19 <https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19>

イ.　（オープン）データセットの共有

政府等が民間企業、大学等と共同で COVID-19 に対するオープンデータを共有する取組が存在する。例えば米国政府は Kaggle 上で COVID-19 Open Research Dataset Challenge[13]を実施している。また EU もデータセットを収容するポータルを設置している[14]。我が国でも内閣官房[15]、自治体等で各種データを公開している。さらに、自治体ごとのデータに対し、AI 等を利用した分析や見える化を統一的に行えるようにするためのデータ項目の定義を行う動きも存在する[16]。

ウ.　計算環境共有

研究者等が AI 等を使った計算を行えるようにするため、計算環境を共有する取組が存在する。海外では、Folding@home[17]のように計算資源に処理（タスク）を分散して配分するプロジェクトや、COVID-19 High Performance Computing Consortium[18]などのように計算環境を無償で提供する取組が存在する。我が国でも、理化学研究所[19]、産業技術総合研究所[20]、高度情報科学技術研究機構[21]等で無償提供の取組が存在する。

エ.　オープンソースによる情報共有

後述する各種アプリケーション等のソースコード等を共有する動きが存在する。例えば、加 DarwinAI 社が胸部 X 線撮影による COVID-19 検出の仕組みをオープンソース[22]で提供している。我が国でも東京都が新型コロナ対策サイトのソース[23]を GitHub 上に公開して

13　Kaggle: COVID-19 Open Research Dataset Challenge
<https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge>

14　EU data portal <https://www.europeandataportal.eu/en/about/european-data-portal>

15　新型コロナウイルス感染症対策関係：全国医療機関の医療体制の状況（G-MIS データ）及びオープンデータを公開しました（β版）
<https://cio.go.jp/hosp_monitoring_c19>

16　新型コロナウイルス感染症対策サイトのためのデータ公開支援
<https://www.code4japan.org/activity/stopcovid1>

17　<https://foldingathome.org/>

18　The COVID-19 High Performance Computing Consortium
<https://covid19-hpc-consortium.org/>

19　理化学研究所：新型コロナウイルス対策を目的とした研究開発を「富岳」上で実施します
<https://www.r-ccs.riken.jp/library/topics/fugaku-coronavirus.html>

20　産業技術総合研究所：AI 向けクラウド型計算システム「ABCI」を新型コロナウイルス感染症対応に無償提供
<https://abci.ai/ja/link/covid-19.html>

21　一般財団法人高度情報科学技術研究機構　「新型コロナウイルス感染症対応の研究」を支援するため HPCI スーパーコンピュータ資源を無償で提供
<https://www.hpci-office.jp/materials/press_20200407.pdf>

22　COVID-Net Open Source Initiative
<https://github.com/lindawangg/COVID-Net/>

23　<https://github.com/tokyo-metropolitan-gov/covid19>

いる。オープンソースとして共有することで、本節**イ**に記載したとおりデータフォーマットの共通化も進むことが期待される。また、マスクやフェイスシールド等、モノ（ハードウェア）のオープンソース化も進められており、リモートでのモノの仕様の共有・共通化に寄与している。

オ． ハッカソン、グランドチャレンジ

COVID-19 対策を題材としたハッカソンやグランドチャレンジについても海外、国内問わず行われている。海外では AI にフォーカスしたものとしてロンドンの CoronaHack - AI vs. Covid-19[24]、国連開発計画が行う The COVID-19 Detect and Protect Challenge[25]、また、前述した COVID-19 Open Research Dataset Challenge や NASA が主催する SPACE APPS COVID-19 CHALLENGE[26]等が存在する。我が国でも異能 vation[27]や未踏 [28]などの取組が存在する。

（２） 検出

ア． 初期警告

AI により感染症の拡大の予兆が存在することを分析する動きが存在した。加・BlueDot 社の AI を利用した初期警告システム [29]は、主なニュース、オンラインコンテンツ、およびその他の情報チャネルを複数の言語でマイニングすることにより疫学的パターンを検出し、初期（2019 年（令和元年）12 月）の段階で感染拡大を警告している。

イ． 診断支援

感染を制限し、病気の広がりを理解するには、迅速な診断が鍵となる。画像と症状データに AI を適用したものは、COVID-19 症例を迅速に診断するのに有用な支援手段の 1 つとなる。前述の加 DarwinAI 社の胸部 X 線撮影による COVID-19 検出の取組がその一例である。

---

[24] CoronaHack - AI vs. Covid-19 - Hackathon in London
<https://www.hackathon.com/event/coronahack---ai-vs-covid-19-99337559314>
[25] <https://www.covid19detectprotect.org/>
COVID-19 検出と保護を可能にするオープンソースハードウェア募集の取組。
[26] NASA: SPACE APPS COVID-19 CHALLENGE
<https://www.spaceappschallenge.org/>
毎年開催しているグランドチャレンジイベントの特別編であり、JAXA もデータを拠出。
[27] 異能 vation
<https://www.inno.go.jp/>
[28] 未踏 第 2 期 AI フロンティアプログラム（After/With COVID-19 対策 AI 活用特別枠）
<https://www.mitou.org/projects/ai_frontier/index_2nd_covid19.html>
[29] BlueDot: Infectious disease surveillance automated and personalized to what's relevant for you.
<https://bluedot.global/products/>

（３） 感染予防

ア． 予測

AI はウイルスの感染の連鎖を特定し、より広範な経済的影響を監視するのに有用である。例えば、ジョンズホプキンス大学は、確認されたコロナウイルスの症例、回復、および死亡に関するライブニュースとリアルタイムデータを通じてウイルスの広がりを追跡するインタラクティブなダッシュボード[30]を利用可能にしている。

また、米・Northeastern 大学では、EpiRisk[31]と呼ばれるツールを開発している[32]。これは、感染した個人が旅行を介して世界の他の地域に病気が拡がる確率を推定するものである。

イ． 人流分析

我が国では、特定エリアの人の増減を見るため、市街地を中心とした人流データの可視化が重要視されている[33]。例えば、内閣官房新型コロナウイルス感染症対策ホームページ[34]でも「人流の減少率」として取り上げられている。

ウ． 接触確認

多くの国・地域等が接触確認アプリを使った感染者接触把握を実施しているが、その方法は接触把握方法（Bluetooth、位置情報他）、陽性者データ管理方法（中央サーバ型、分散型等）により様々である[35]。また、その目的も、接触度に応じた制限・隔離、公衆衛生当局に

---

[30] <https://github.com/CSSEGISandData/COVID-19>
（ダッシュボード）
<https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>

[31] <http://epirisk.net/>

[32] 米 Northeastern 大学：EpiRisk
<https://www.gleamproject.org/tools/epi-risk>

[33] 新型コロナ対策に生きる「統計データ」、提供する各社の違いを一挙公開
<https://xtech.nikkei.com/atcl/nxt/column/18/01304/051900003/>

[34] 内閣官房：「人流の減少率」
<https://corona.go.jp/#area-transition>

[35] AI とは直接関係しないが、これらの動きに対し、Google、Apple 両社は、スマートフォンによる同アプリを作成するための Exposure Notification API を公表しており、2020 年（令和 2 年）5 月 20 日現在、22 カ国及び米国のいくつかの州で使用が計画されているとの報告がされている。
Apple 社発表
<https://developer.apple.com/documentation/exposurenotification/building_an_app_to_notify_users_of_covid-19_exposure>
Google 公式ブログ
<https://blog.google/inside-google/company-announcements/apple-google-exposure-notification-api-launches/>
また、わが国でも以下のとおり、同年 6 月、上記 API を活用した接触確認アプリを公表している：
新型コロナウイルス接触確認アプリ（COCOA) COVID-19 Contact-Confirming Application

よる濃厚接触者把握のための補完、通知を受けたユーザの行動変容による感染拡大防止（当局も把握せず）など様々存在する[36]。なお、同アプリが有効に動作するためにはその導入が増えることがポイントと言われており、そのための手段も重要となる。特に、セキュリティ・プライバシー等への配慮の必要性についての議論[37]が各地で起こっており、法案を起草・提出する動きもある[38]。

我が国でも、こういった動きに対し、国民・市民の安全とセキュリティ・プライバシーの確保を両立する観点等から、様々な議論が起こっている[39]。

エ. 誤／偽情報[40]対策

誤情報・偽情報の拡散（COVID-19 の「インフォデミック」[41]）は大きな問題となっており[42]、その対策のため、Twitter[43]、Facebook[44]などの各種 SNS では AI 等を活用し、プラ

（厚生労働省ホームページ）
< https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/cocoa_00138.html >

[36] 新型コロナウイルス感染症対策 テックチーム：接触確認アプリの導入に係る 各国の動向等について（第 3 回資料 1-2）
<https://cio.go.jp/sites/default/files/uploads/documents/techteam_20200508_02.pdf>

[37] 例えば、Coronavirus: a common approach for safe and efficient mobile tracing apps across the EU：
<https://ec.europa.eu/digital-single-market/en/news/coronavirus-common-approach-safe-and-efficient-mobile-tracing-apps-across-eu>

[38] 例えば、COVID-19 Consumer Protection Data Act of 2020
<https://www.commerce.senate.gov/services/files/A377AEEB-464E-4D5E-BFB8-11003149B6E0>

[39] 例えば、以下等：
岸本・工藤（大阪大学社会技術共創研究センター）「接触確認アプリと ELSI に関する 10 の視点と 3 の提言 Ver.0.9」
<https://elsi.osaka-u.ac.jp/research_category/elsi_note/>
まもりあい note（Civic tech による接触確認アプリ「まもりあい Japan」検討を通じたアプリのあり方等の検討状況に言及。）
<https://note.com/hal_sk/m/m53cefeea1340>

[40] ここでは、総務省「プラットフォームサービスに関する研究会　最終報告書」（2020 年（令和 2 年）2 月公表）に従い、「誤情報」とは単なる誤った情報のこと、「偽情報」とは何らかの意図性を持った虚偽の情報のことをそれぞれ指すものとする。同報告書は次に掲げる URL のウェブサイトに所掲。
< https://www.soumu.go.jp/main_content/000668595.pdf >

[41] Information(情報) + epidemic (伝染)からなる造語。SNS 等を通じ、不確かな情報が拡散される現象を指す。

[42] 総務省では、COVID-19 の「インフォデミック」が問題となっている実態について、以下のとおり報告している。:
「新型コロナウイルス感染症に関する情報流通調査」
< https://www.soumu.go.jp/menu_news/s-news/01kiban18_01000082.html >

[43] Twitter: Coronavirus: Staying safe and informed on Twitter
<https://blog.twitter.com/en_us/topics/company/2020/covid-19.html>

[44] Facebook: An Update on Our Work to Keep People Informed and Limit Misinformation About COVID-19
<https://about.fb.com/news/2020/04/covid-19-misinfo-update/>

ットフォーム上の問題のある素材を検出し削除するよう努めている。**第６章４．（２）**とも関連するため参照されたい。

（４） 感染者等の対処

ア． （ロボット等による）配達

病院などの緊急のニーズに応えるために、ロボット等を活用した様々な配達が行われている。例えば中国ではドローン・ロボット等を使った食品や医薬品の配送、空中散布・消毒などが行われている[45]。また、その他の国でも、医療従事者の人手不足を補ったり、院内感染を防いだりするために、病院内での消毒作業や病院入口での患者のスクリーニング等にロボットを活用する取組が進められている[46]。

イ． サービス

医療機関のサポートを目的としたバーチャルアシスタント、チャットボットの導入が世界各国で進んでいる。これらのツールは、症状の存在に応じた選別に役立っている。例えば米国疾病予防管理センターとマイクロソフトはコロナウイルスセルフチェッカーサービス[47]を開発し、ユーザーが COVID-19 を自己評価して一連の行動を提案できるようにしている。

（５） オンライン化（リモート作業）への対応

ア． オンライン教育

リモートでの教育推進の一環として、学生等が各々の学習の進捗にあわせて適応的に授業を進められるアダプティブラーニングを導入する取組が進められている[48]。

イ． 業務自動化等

テレワーク等を促進するための手段の 1 つとして、RPA やチャットボットの活用など各種業務を自動化する取組が進められている。

---

[45] 3 ways China is using drones to fight coronavirus
<https://www.weforum.org/agenda/2020/03/three-ways-china-is-using-drones-to-fight-coronavirus/>
[46] 薬の運搬や病室の殺菌を肩代わり ウイルス感染症と戦うロボット
<https://project.nikkeibp.co.jp/mirakoto/atcl/robotics/h_vol35/>
[47] Testing for COVID-19
<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/testing.html>
[48] アダプティブラーニングとは？主要な学習ツール、メリット、デメリットを徹底解説！
<https://coeteco.jp/articles/10634>

（６） 経済等の回復

ア． 人流分析

AIによる分析で経済回復をモニタリングする動きがある。例えば、中国WeBank社では、衛星写真、SNS およびその他のデータ（Google のコミュニティモビリティレポートなど）を分析することにより、経済危機と回復のモニタリングに有用とされている[49]。

また、後述する ABEJA 社は、AI を使った同社小売店解析システムを使った分析により COVID-19 による売上・客数への影響調査を発表している[50]。

こういった分析を通じて、（経済回復の指標とするのに加え）将来の感染（拡大）の発生に対する早期警告を作成できる可能性がある。

２． 国内・海外及び国際的な議論の動向

（１） 文部科学省：信頼される AI

文部科学省では、「AI 戦略 2019」[51]に記載のある「信頼される高品質な AI」（ Trusted Quality AI ）の開発の重要性を踏まえ、今後の AI の進化と信頼性確保のための基盤技術の開発を行うことを 2020 年度（令和 2 年度）の戦略目標として掲げることを決定し、同年 3 月 9 日に公表した[52]。

現在の AI 技術の中心である深層学習（ディープラーニング）においては、様々な対策が喫緊の課題となっており、産業界ではコンソーシアム等による対策の検討も始まっているが、その限界を超えた AI 技術そのものの発展・革新が必要であるほか、社会からの要請に応え得る根本的な信頼性確保が求められることから、本戦略目標では「人間中心の AI 社会原則」に基づいた「信頼される高品質な AI」（Trusted Quality AI）の創出に向けた研究開発を推進することとしている。具体的には、以下の 3 つの達成を目指している：

- 現在の AI 技術の限界を克服する新技術の創出
- AI システムの信頼性・安全性を確保する技術の創出
- データの信頼性確保及び意思決定・合意形成支援技術の創出

---

[49] IEEE: Satellites and AI Monitor Chinese Economy's Reaction to Coronavirus <https://spectrum.ieee.org/view-from-the-valley/artificial-intelligence/machine-learning/satellites-and-ai-monitor-chinese-economys-reaction-to-coronavirus>

[50] 【ABEJA Insight for Retail】アパレル・雑貨店の来店者数、3 月 2 週から回復の兆し？ 新型コロナ影響の検証第 2 弾 <https://abejainc.com/ja/news/article/20200325-2681>

[51] 「AI 戦略 2019」（2019 年（令和元年）6 月統合イノベーション戦略推進会議決定）〈https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistratagy2019.pdf〉

[52] 次に掲げる URL のウェブサイトに所掲。<https://www.mext.go.jp/b_menu/houdou/2020/mext_00487.html>

（２） 欧州連合（EU）

EU の行政執行機関である欧州委員会は 2020 年（令和 2 年）2 月 19 日、欧州のデジタル未来形成のため、「AI 白書」を公表した[53]。AI の普及を促進しつつ、AI の信頼できる安全な開発を可能にするための政策オプションを卓越性（excellence）と信頼性（trust）の観点から提示している。

卓越性の観点では、バリューチェーン全体にリソースを動員し、中小企業等の AI の展開を加速するためのインセンティブの作成に言及しており、加盟国との連携、研究・イノベーションコミュニティの取組強化（テストセンター構築など）、スキル、中小企業、民間部門との連携について言及している。

他方、信頼性の観点では、リスクの低いシステムに過度の負担をかけることなく、リスクの高い AI システムに対処できるよう、リスクの高い領域・用途等のスコープを明確にした上で、リスクの高低に応じた将来の規制の在り方を提示している。また、消費者保護、不公正な商慣行に対処し、個人データとプライバシーを保護するための厳格な EU 規則の適用を継続することに言及している。

リスクが高いことが想定されるヘルスケア、輸送、公共部門等の特定の用途においては、AI システムは透明で追跡可能であり、人間の監視を保証すること等、EU の信頼性のある AI 倫理ガイドラインに記載の条件等を踏まえることが必要としている（他、適切に機能するためのトレーニング、偏りのないデータの利用など）。特に、遠隔生体認証のための顔認識の使用について、現在 EU では特定の条件を除き一般的に禁止されており、EU または国内法に基づいて例外として正当化された場合にのみ使用できるが、これについて幅広い議論を開始するとしている。

また、リスクが低い AI システムにおいても、信頼を醸成すべく、EU における客観的なベンチマークによる任意のラベル付けスキーム、すなわち認定等の仕組みを検討するとしている。

なお、同案について同年 5 月 19 日まで[54]意見募集を実施することとしている。また、本意見募集に対し、本推進会議及び同ガバナンス検討会構成員有志から意見が提出されている[55]。

また、EU 理事会は同年 6 月 9 日、上記の「AI 白書」を含めたデジタル戦略を支持する内容を公表した[56]。

---

[53] “WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust”
次に掲げる URL のウェブサイトに所掲。
<https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf>

[54] 締め切りは 6 月 14 日に延長された。

[55] 提出意見は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

[56] 次に掲げる URL のウェブサイトに所掲。

（３） 米国

米国行政管理予算局（OMB）は、2020 年（令和 2 年）1 月 13 日、「省庁が民間の AI 導入に係る規則策定時に従うべき覚書案」に対する意見募集を開始した[57]。

本覚書案は、2019 年（平成 31 年）2 月に出された米大統領令（AI イニシアチブ）[58]を受け、各省庁が民間の AI 導入に係わる規則を策定する際に従うべき原則を記したものである。

原則は 1) AI に対する一般市民の信頼、2) 規則作成手続きへの一般市民の参加、3) 科学的公正性と情報品質、4) リスク評価・管理、5) 費用対効果分析、6) 柔軟性、7) 公平性・非差別性、8) 情報開示と透明性、9)安全とセキュリティ、10) 省庁間連携の 10 項目から構成されている。意見募集は 2020 年（令和 2 年）3 月 13 日まで行われ、全部で 81 の意見があったとされている。

前述の大統領令によると、本意見募集の後、正規の覚書が OMB から発出され、その後、180 日以内に各省庁の長は、覚書との一貫性を確保するための実施計画を提出することになっている。

また、米国科学技術政策局（OSTP）は、2020 年（令和 2 年）2 月、前述の大統領令を受けた年次報告書を公表しており、AI の研究開発への投資、AI のリソース開放、AI のイノベーション障壁の除去、AI に対応できる労働力の向上、米国の AI イノベーションを支える国際環境の促進及び政府のサービス及びミッションのための信頼できる AI の採用の視点でまとめている[59]。

また、米国国防総省（DoD）は 2020 年（令和 2 年）2 月 24 日、倫理的に AI を導入するための原則を正式に採用した[60]。本原則は、同省の国防イノベーション委員会が 2019 年（令和元年）10 月に提示した内容に基づいており、①AI の開発・導入・利用に責任を有す

< https://www.consilium.europa.eu/en/press/press-releases/2020/06/09/shaping-europe-s-digital-future-council-adopts-conclusions/ >

[57] Draft to the Heads of Executive Departments and Agencies, “Guidance for Regulation of Artificial Intelligence Applications”
次に掲げる URL のウェブサイトに所掲。
<https://www.federalregister.gov/documents/2020/01/13/2020-00261/request-for-comments-on-a-draft-memorandum-to-the-heads-of-executive-departments-and-agencies>

[58] “Executive Order on Maintaining American Leadership in Artificial Intelligence”
次に掲げる URL のウェブサイトに所掲。
<https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>

[59] AMERICAN ARTIFICIAL INTELLIGENCE INITIATIVE: YEAR ONE ANNUAL REPORT
次に掲げる URL のウェブサイトに所掲。
<https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>

[60] DOD Adopts Ethical Principles for Artificial Intelligence
次に掲げる URL のウェブサイトに所掲。
<https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/>

ること、②偏向を最小限に抑えること、③AI の技術等について透明であること、④安全・セキュリティ対策がとられていること、及び⑤意図しない動作を完全停止できる統治可能性を有していることの 5 つを提示。さらに実際に運用していくためのガイダンスが必要であると述べている。

（４） 経済協力開発機構（OECD）

経済協力開発機構（OECD）は、2019 年（令和元年）11 月 21 日から 22 日にかけてデジタル経済政策委員会（CDEP）会合を開催した。AI については、同年 5 月に策定された AI に関する理事会勧告の履行に係るプラクティカル・ガイダンス案についての議論が行われるとともに、AI 政策に関するオブザーバトリ（OECD.AI）のデモンストレーション、非公式の専門家ネットワーク（後述）の目的や運用方針及び OECD.AI にパートナーとして参画・貢献する者に係る基準（案）について事務局から説明がなされた。また、フランス・カナダより、Global Partnership on AI（GPAI）（後述）についての説明がなされた。

さらに OECD は、2020 年（令和 2 年）2 月 26 日から 27 日にかけて、AI に関する専門家ネットワーク（ONE AI: OECD Network of Experts on AI）の第 1 回会合を開催。ONE AI は OECD が AI に関する分析作業を進めるために専門家の意見を提供する非公式の諮問グループで、マルチステークホルダかつ学際的な専門家から構成[61]されており、第 1 回会合においては、特に AI の分類手法、及び人間中心の AI の実現に向けたイニシアチブの紹介が行われた。

また、OECD は前述の ONE AI 会合に引き続き、同 27 日に OECD.AI[62]のローンチ（運用開始）イベントを開催した。OECD.AI は OECD 関連委員会及び OECD 以外の政策立案者が、AI 政策の課題、解決策及び計測方法に対応し、AI に関する取組の情報共有を進めるためのプラットフォーム（ライブ型データベース）で、以下の 4 つの柱の活動から構成されており、AI に関する情報共有をはじめ、政策機会の活用や課題の解決手段を提供するものである。

- **AI 原則**：OECD の AI 原則及び実務者向けのガイダンス（プラクティガルガイダンス）を掲載。
- **政策分野**：公共政策分野毎に、AI 政策ニュースや AI 調査に関する公表内容等の様々なコンテンツにアクセス可能。
- **トレンドとデータ**：AI に関する調査データを掲載。 データの地域比較や時間的変化を観ることが可能。
- **国々とその取組**：AI に関する国家戦略や政策、取組に関するデータベースであり、各国の AI 政策を共有・比較することが可能。

内容は今後も継続して更新される予定である。

---

61 我が国からは、本推進会議の須藤議長が選出されている。

62 <https://oecd.ai/>

（５） GPAI

2020 年（令和 2 年）6 月 15 日、「人間中心」の考えに基づく責任ある AI の開発と使用に取り組む国際的なイニシアチブである「AI に関するグローバルパートナーシップ（Global Partnership on AI、GPAI）」が設立され、我が国を含むG7 各国及びオーストラリア、インド、メキシコ、ニュージーランド、韓国、シンガポール、スロベニア、欧州連合（正式手続中）による共同宣言が行われた[63]。同共同宣言には以下のように記載されている：

設立メンバーとして、我々は、OECD による AI 勧告に詳述されているように、人権、基本的自由、そして我々が共有する民主主義の価値観に調和した形で、責任ある、人間を中心とした AI の開発と利用を支持する。そのため、我々は他の関心を有する国やパートナーとの協働を期待する。GPAI は、AI の責任ある開発と利用を導く、国際的で多様な関係者が参画する取組みであり、人権、包摂性、多様性、イノベーション、そして経済発展を土台とするものである。その目指すものを達成するために、この取組みでは、AI に関する重要事項について、最先端の研究と実装を支援することにより、AI に関する理論と実践の隔たりに橋渡しをすることを目指している。パートナーや国際機関との協力により、GPAI は、以下の 4 テーマの作業部会でともに作業を進める一流の専門家を、産業界、市民団体、政府、そして学術界から呼び集める。

1) 責任ある AI
2) データガバナンス
3) 仕事の未来
4) イノベーションと商業化

短期的な重要事項として、GPAI の専門家は、COVID-19 へのより良い対応とそこからの回復に AI をいかに活用できるかについても、検討する。

また、本共同宣言に先立ち、同年 5 月 28 日、G7 科学技術大臣会合がオンラインで開催され、COVID-19 に関する大臣宣言[64]が採択された。同宣言の中では GPAI の設立について以下のように記載されている：

COVID-19 への対応及び COVID-19 からの回復などを当初の焦点として、我々の共通の民主的な価値観を反映し、共通のグローバルな課題に対応する AI の発展におけるマルチステークホルダー協力を促進するため、カナダとフランスが議長国を務めた 2018 年と 2019 年の G7 で構想された、**AI に関するグローバル・パートナーシップ（GPAI）を立ち上げる。**人権、基本的自由、我々の共通の民主的な価値観と一致した形で、責任ある、また人間中心の AI の開発及び利用を約束する。

63 次に掲げる URL のウェブサイトに所掲。
< https://www.soumu.go.jp/menu_news/s-news/01tsushin06_02000204.html >
64 次に掲げる URL のウェブサイトに所掲。
< https://www.state.gov/g7-science-and-technology-ministers-declaration-on-covid-19/ >

（６） ユネスコ（UNESCO）

ユネスコ（UNESCO）は 2019 年（令和元年）11 月に開催された第 40 回ユネスコ総会において、AI 倫理に関するグローバルな勧告を 2021 年の総会を目処に策定することを決議し、その作成作業を開始した。

また、ユネスコは、同勧告の草案を作成するために、2020 年（令和 2 年）3 月 10 日、AI の倫理についてのアドホック専門家グループ（AHEG）を結成するために専門知識を持つ 24 名の有識者を任命した[65]。同グループの第 1 回会合が同年 4 月 20 日から 24 日にかけてオンライン形式で開催され、草案（初版）に関する議論が行われた[66]。同会合においてアズレー事務総長は「COVID-19 により様々なデジタル技術の利用が増加している。これは AI の開発に関する既存の倫理課題を浮き彫りにした。したがって、AHEG が規範文書の草案の作成を開始することは重要だった」と述べている。

今後は、同年 7 月まで複数のステークホルダと協議を重ねていく予定である。

（７） G20 デジタル経済大臣会合

2020 年（令和 2 年）の議長国であるサウジアラビアは、同年 4 月 30 日、G20 デジタル経済大臣臨時会合をテレビ会議形式で開催し、G20 デジタル経済大臣臨時会合 COVID-19 への対応声明を採択した[67]。

同声明の中で、「安全な方策によるデータの交換」と題し以下の記述がある：

> 「COVID-19 に関する不確実性と、パターン認識を加速し、証拠に基づく政策立案を可能とするデータと人工知能（AI）の力を認識し、我々は、COVID-19 及び他の感染症の更なる拡大の監視と理解、予防に寄与し得る、信頼性があり正確な非個人情報を収集、蓄積、処理、共有するための協力を奨励する。 COVID-19 に関連するデータは、国際保健規則(IHR) 2005 及び国内法令に従って、個人のプライバシーとデータセキュリティを保護する倫理的、透明、安全、相互運用可能、かつ安心な方法で共有され、処理されるべきである。我々は、データまたはアルゴリズムにおける潜在的なバイアスが適切に対処されることを確保する必要性を認識する。」

また、「健康分野におけるデジタル技術の研究開発」と題し、パンデミックに対する闘い及び予防に貢献する、AI を含むデジタル技術の可能性を認識するとともに、同技術の研究への投資の増加について述べている。

---

65 我が国からは、本推進会議の須藤議長が選出されている。

66 次に掲げる URL のウェブサイトに所掲。
<https://en.unesco.org/news/unescos-international-expert-group-begins-work-drafting-first-global-recommendation-ethics-ai>

67 次に掲げる URL のウェブサイトに所掲。
<https://www.soumu.go.jp/menu_news/s-news/01tsushin06_02000202.html>

第２章 AI ネットワーク化の進展に伴い形成されるエコシステムの展望

１． 背景及び分析方針

（１） 背景

本推進会議では『報告書 2018』において、AI の利活用に関するガバナンスの在り方に関する検討に当たり、その前提として AI ネットワーク化の進展に伴い形成されるエコシステムの展望（以下「エコシステムの展望」）を行った。

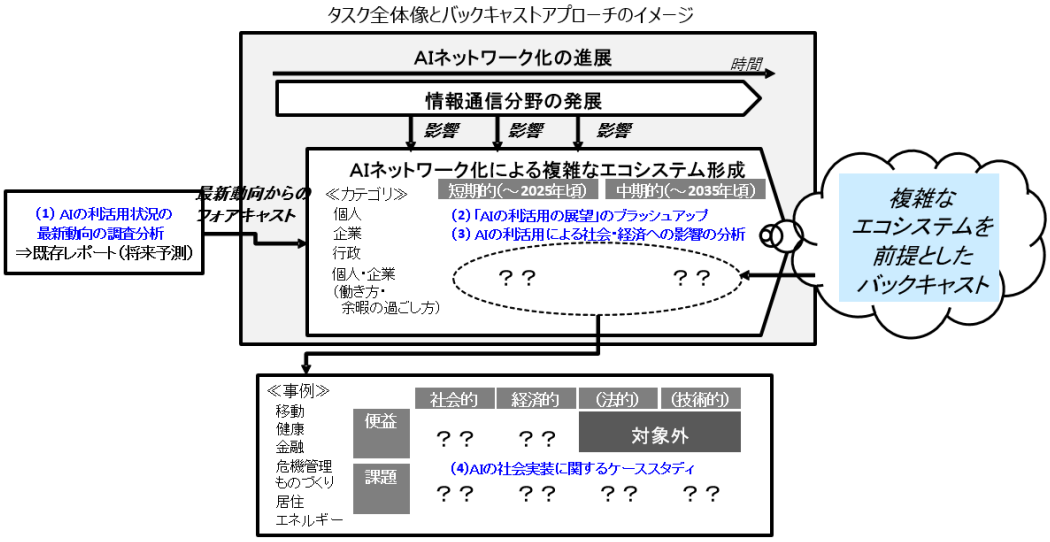
この展望の結果を踏まえ、AI の利活用に関する便益及びリスクの分析を行い、各分野に共通する課題を踏まえ、AI 利活用原則を含む AI 利活用ガイドラインの策定を行ってきた。

このような中で、『報告書 2018』で示した検討結果から技術及びそれに基づく利活用が急速に進展しているため、特に同報告書公表以降に出された各種情報を踏まえ、その内容について更新した。

なお、AI・ビッグデータ・IoT 等に代表される急速な ICT の進展により、産業構造の変革期を迎え、人々の暮らしも変容していく社会において、社会全体のエコシステムのありようを俯瞰・想定した中で「AI の利活用の展望」を検討する必要がある。

また、AI ネットワーク化が健全に進展していくための方策を検討するという目的に鑑みた場合、「AI の利活用の展望」は下図のとおり、将来的にエコシステムが複雑化していくことを前提として、様々な背景から（フォアキャストのみならずバックキャストを含めて）抽出された利用シーンを基に便益や課題を評価する必要があると想定する。

タスク全体像とバックキャストアプローチのイメージ



（２） 分析方針

ア． AI の利活用シーンの展望

本展望を行うにあたっては、『報告書 2018』同様、生活者と事業者の両面から、AI を利活用する**エンドユーザー視点**での領域整理を行った。利用シーンについては、『報告書 2018』で対象とした「個人の利用」及び「企業の利用」に「行政の利用」の観点を加えた。

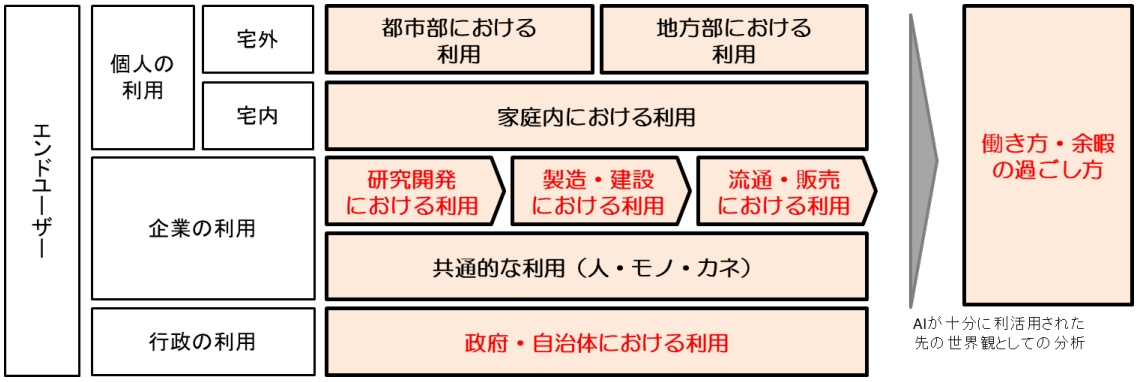
「個人の利用」については、『報告書 2018』同様、「宅外」と「宅内（家庭内における利用）」に分類した上で、「宅外」はさらに「都市部における利用」と「地方部における利用」に分割した。

「企業の利用」についても、『報告書 2018』同様、「共通的な利用」に関する分析を実施した。さらに、企業内でのプロセスに着目し、「研究開発における利用」「製造・建設における利用」「流通・販売における利用」に分割し分析を実施した。

また、これらの領域で十分に AI が利活用された場合に、働き方や余暇の量が大きく変わる可能性があることから、AI の利用シーンとしての領域ではなく、AI が十分に利活用された先の世界観を分析するために「働き方・余暇の過ごし方」の領域を定義し、分析を行った。

具体的には下図のとおりである。

エンドユーザー視点での領域整理



（『報告書 2018』 別紙 2「AI ネットワーク化の進展に伴い形成されるエコシステムの展望について」をもとに作成。赤字は『報告書 2018』に対し加えた領域）

また、COVID-19 感染拡大防止の一環で、リモートワークの推進など新しい日常（ニューノーマル）が期待されており、AI を利活用する者のライフスタイルもそれに併せて変化している。COVID-19 への対応に伴う AI の利活用の事例については**第１章１．**でも述べているが、本章でも触れる。

イ． AI の社会実装に関するケーススタディ

上記利活用シーンをもとに、次に掲げる 7 つの事例に関するケーススタディを行い、AI の利活用による便益及び課題を整理した。

○ 金融
○ ものづくり
○ 居住
○ 健康（医療・介護）
○ 危機管理（防犯・公共インフラ・防災）
○ 移動（完全自動運転）
○ エネルギー

２． AI の利活用の展望

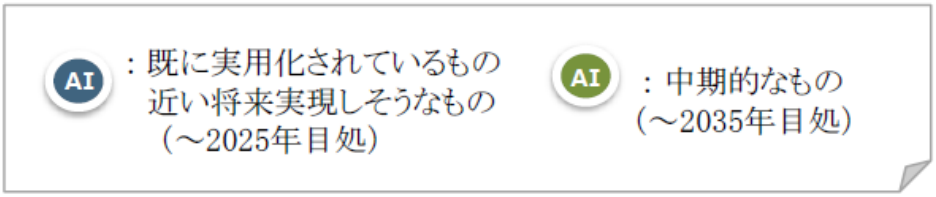
本章１．に記載のとおりAIの利活用シーンを分類した上で、次に掲げるように、それぞれの分類において主な利用シーンを想定してAIの利活用を展望した。

利用場面			利用シーン
個人の利用	宅内	都市部	移動、介護、 観光・旅行、教育・人材育成
		地方部	移動、医療、 仕事、生活環境
	宅外	家庭内	ヘルスケア、家事、 安全・快適な居住環境、ライフスタイル
企業の利用		研究開発	研究、開発、 共通（効率化）、共通（知見の集約）
		製造・建設	設計、生産計画、 製造、調達
		流通・販売	物流、広告、 アフターフォロー、販売
		共通的（な利用）	ヒト、モノ（オフィス環境）、 カネ、情報
行政の利用		政府・自治体	政策立案（政府）、行政事務・執行（政府） 政策立案（都道府県・基礎自治体）、 行政事務・執行（都道府県・基礎自治体）
働き方・余暇の過ごし方			働き方、余暇の過ごし方

これらの利用の場面について、次に掲げる都市部における利活用の例のように、AI の利活用を展望した。それぞれの利用の場面における AI の利活用の展望の詳細は、**【別紙２－１】**のとおりである[68]。なお、分析にあたり、以下の点に留意した。

- **利用シーンの実現時期**

利用シーンを分析する際には、実現時期について、「既に実用化されているもの/近い将来実現しそうなもの（～2025 年目処）」と、「中期的なもの（～2035 年目処）」に分類した。分類を実施する際には、サービスの普及率や機能の充実度を勘案せず、サービスとして上市されるタイミングを実現時期とした。



- **企業における「サービス提供時の AI 利用」の位置づけ**

本展望では、エンドユーザー視点で領域を整理しているため、企業がサービス提供する際の AI 利活用については、個人向けのサービスであれば個人の利用シーンに包含され、サービスを提供するために企業内で利用される場合にはプロセスや共通に包含されるため、企業の分類において「サービス提供」の領域を定義していない。

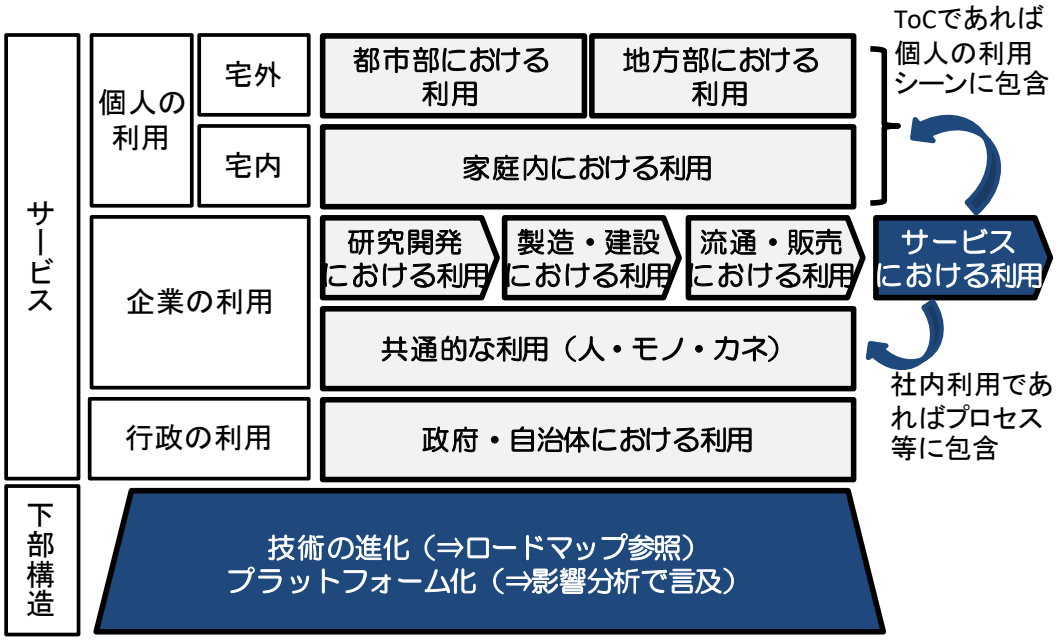
- **下部構造の分析**

本展望では、技術の進化やビジネスモデルの変化（プラットフォームビジネス等）を、AI の利用シーンの下部構造と捉えている。このため、技術の進化については以下ロードマップ等を参考に利用シーンを抽出することで考慮し、ビジネスモデルの変化については、利用シーンとしてではなく、影響分析の部分で言及することとした。

---

68 ここでは、現時点における法律等及び実用化されている技術や研究が進められている技術の水準を前提としたときに実現が困難であると見込まれるものであっても、将来的な利活用の可能性を展望して記載している。また、実用化にあたっては、経済的なコスト等を勘案することとなる点に留意することが必要である。

エコシステムの展望の分析における留意点



参考にしたロードマップ（例）

分類		発表元	ロードマップ・戦略等	発表日
領域全般		内閣府	AI戦略 2019	2019年6月
		NEDO	人工知能の研究開発目標と産業化のロードマップ	2017年3月
		総務省	IoT新時代の未来づくり検討委員会　未来をつかむTech戦略	2018年8月
		総務省	新たな情報通信技術戦略の在り方 第3次中間答申	2017年7月
		総務省	新たな情報通信技術戦略の在り方 第2次中間答申	2016年7月
		総務省	地域IoT実装推進ロードマップ（改定）	2018年4月
		文部科学省	科学技術発展による社会の未来像	2019年11月
個別領域	行政	総務省	地方自治体における業務プロセス・システムの標準化及び AI・ロボティクスの活用に関する研究会（スマート自治体研究会）	2019年5月
	医療・介護	厚生労働省	保健医療分野AI開発加速コンソーシアム資料 参考資料3 「保健医療分野AI開発加速コンソーシアム議論の整理と今後の方向性」	2019年3月
	ヘルスケア	経済産業省	2040年における未来の医療・福祉・介護分野の在り方と ロードマップ策定等に関する調査	2019年10月
	移動	内閣府	官民ITS構想・ロードマップ2019	2019年6月
	製造	経済産業省	スマートファクトリーロードマップ	2017年5月
	建築	国土交通省	AI開発支援プラットフォームの開設準備ワーキング・グループについて	2019年8月

３．　AI の社会実装に関するケーススタディ

本章２．の AI の利用シーン等をもとに、本章１．に述べた７つの事例（金融、ものづくり、居住、健康（医療・介護）、危機管理（防犯・公共インフラ・防災）、移動（完全自動運転）、エネルギー）に関するケーススタディを行い、AI の利活用により想定される便益及び課題のうちいくつかの例を整理した。

特に課題の例については、下図のとおり、社会的、経済的、技術的、法的課題についてそれぞれ挙げるとともに、AI が導入される前に考えられる課題と導入された後の課題をそれぞれ整理した。

各ケーススタディの詳細は、**【別紙２－２】**のとおりである[69]。

ケース：移動（完全自動運転）

想定される便益（例）

○ 人間は運転する必要がなくなり、自動車での移動において、移動時間を有効に活用することができるようになる。

○ 高齢者や障害者の方にとって、手軽に移動することができる手段が確保されることになり、病院や買物などに容易に出かけることができるようになる。

○ 深夜や早朝などにおける長距離トラックや長距離バスの運転をする必要がなく、働き方やワーク・ライフ・バランスを見直すことなどができるようになる。

○ 特に地方部などにおける路線バスの運転手不足などの問題を改善することができ、路線の廃止・縮小を回避することができるようになる。

想定される課題（例）

	実現前	実現後
社会	• 技術的に安全性が担保されているか、事故が起きた場合に誰が責任を負うかという問題が不透明であることから、自動運転に抵抗感を感じ、サービスが受け入れられない可能性がある。	• これまで通勤、通学等に充てていた時間を他の目的のために利用する流れが生まれることにより、個人の住む場所や生活スタイルにこれまでとは大きな変化が生じる可能性がある。
経済	• インフラ協調型の自動運転の場合、財政がひっ迫している地方自治体では、インフラが整備されず、自動運転の普及に関して地域間格差が生じる可能性がある。 • 配送や輸送サービスに関わる企業や従業員の職が減少することにより、代わりに従事する職が見つけにくくなり、導入に対して後ろ向きになる可能性がある。	• 普及が先行している自動車メーカーは、販売後の学習データを多く確保できるため、後発の自動車メーカーの参入を阻害する可能性がある。 • 人間が運転しなくなることによって、これまで人的ミスの結果として起きていた事故が大幅に減り、自動車に関係する保険の料金設定を大幅に変えなければならなくなる可能性がある。
技術	• 未学習データに対してAIが予測不能な挙動をする可能性や、データにより正しく学習されていたとしても誤認識・未認識が避けられない可能性がある。	• 市場投入後の世の中の変化にAIが対応できなくなる可能性がある。 • 車間で交渉・調整が成立せずに適切な動作ができない可能性がある。 • AIシステムがハッキング等された場合、その自動運転車が正常に機能しなくなるだけではなく、ネットワークを介して、次々と他の自動運転車にも影響が及び、事故や交通の混乱が生ずる可能性がある。
法律	• AIのブラックボックス化により、自動運転の法的責任の所在を確定することが難しく、自動車メーカーや利用者との間で合意形成が困難になる可能性がある。	• 国内での法整備に加えて諸外国との調整が必要となり、現行の法制度だけでは対応することができなくなるため、自動車メーカーごとに新たな対応を迫られる可能性がある。

（注）想定される便益及び課題のうち、いくつかの例を記載

なお、ここに掲げる想定される課題（例）は、以下のような意味で書かれたものではない：

- （「実現前」の記載について）全て解決しないと何も実現できない。
- （「実現後」の記載について）全て解決しないと実施すべきでない。

むしろこうした課題はビジネス開発や研究開発のチャンスを意味するものであり、今後、企業等の努力によって解決することができれば大きな利益を生むものと考えられる。こういった課題を一つ一つ解決していくこと及び対応していくことが、それを行った主体の付加価値につながるという趣旨で書いているものであり、本推進会議が 2019 年（令和元年）8 月に公表した「AI 利活用ガイドライン」(3 ページ) の以下の記載とも符合するものである。

> AI サービスプロバイダやビジネス利用者がこうした自主的な取組を実施することで、提供する AI サービスや AI を活用した業務に対して付加価値を与えることも可能となりうる。

[69] 今般、報告書 2020 **第２章【別紙２－２】**で取りまとめた各分野の社会実装に関するケーススタディについては、今後、個別のヒアリングを通じて、社会実装を進めていくための途中段階を丁寧に分析し、そこで生じる多くの個別課題に取り組む必要があるのではないか、また、個別課題への取組方についてもあわせて議論する必要があるのではないか、との意見があった。

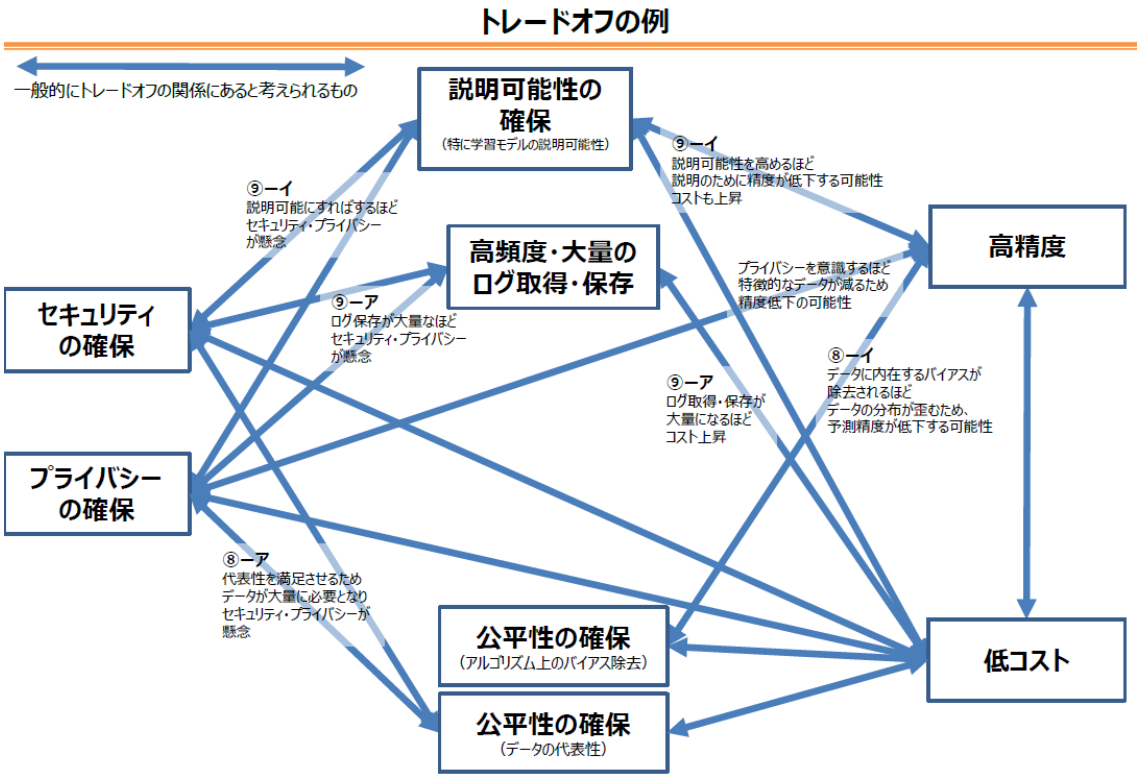
第３章 開発者・AI サービスプロバイダーにおける取組

１． 議論の出発点

開発者・AI サービスプロバイダーにおいては、AI 原則等の策定の取組が進められてきているものの、その取組が十分な広がりをみせている状況ではないと考えられる [70]。もとより、開発者・AI サービスプロバイダーによる AI 原則等の策定は「安心・安全で信頼性のある AI の社会実装」を推進するための自主的な取組の一つにすぎないともいえる。しかしながら、AI の開発、サービスの提供にあたり AI 原則等がどのような意義を有するかについてお互いに共有し理解を深めることにより、AI 原則等の策定の取組が広がり、「安心・安全で信頼性のある AI の社会実装」が進められ易くなるものと考えられる。本推進会議で策定した両ガイドラインもこうした自主的な AI 原則等の策定の取組の参考となる趣旨のものであることは「**はじめに**」でも述べたとおりである。また、AI 原則等は、その具体的な機能として、AI の開発、サービス提供にあたり下図にあるようなトレードオフにある考慮要素を踏まえた事業判断ツールの一つとしての役割を担うことが考えられる。とすれば、AI 原則等の策定を含め、広く AI の開発・利活用に必要なガバナンスにはどのようなものが考えられるかについて理解を深めることも「安心・安全で信頼性のある AI の社会実装」の推進に必要な取組と考えられる。こうした問題意識のもと、『報告書 2019』の「**第２章　ＡＩ利活用ガイドライン策定の考え方　４．今後の展開**」及び「**第３章　今後の課題**」の関連する記述も踏まえ、「**はじめに**」の「**（１）ア及びイ**」の論点を提示し、意欲的な取組を行っている企業からこれらの論点を中心にヒアリングを行うこととした。

---

[70] 国内外における AI 原則等策定の取組状況を精力的に調査したものとして、総務省情報通信政策研究所情報通信法学研究会 AI 分科会 2019 年度（令和元年度）第 1 回会合における新保史生構成員（慶應義塾大学総合政策学部教授）発表資料「AI 原則は機能するか？」https://www.soumu.go.jp/main_content/000660996.pdf　6 頁以下参照。



『報告書 2019』別紙 1（附属資料）　「AI 利活用原則の各論点に対する詳説」より引用

２．　（株）ABEJA（「Ethical Approach to AI（EAA）における取組み」）

（１）　ヒアリング概要 [71]

ア．　サービス概要

ディープラーニングの実装、運用プロセスを効率化する ABEJA Platform 及び顧客行動データの取得・分析を基軸とする AI を活用した店舗解析サービスである ABEJA Insight for Retail を提供。

イ．　ガバナンス体制としての Ethical Approach to AI (EAA) の設置

ABEJA Insight for Retail では、小売りの店舗内にカメラを設置することで、顧客の顔画像を撮ることで来店者の動線分析やリピート分析等を実施。このように顧客の顔画像を取得することから、個人情報保護法・カメラ画像利活用ガイドブック ver2.0[72]等を踏まえ、利用目的の通知等、法令遵守の取組を行ってきているが、プライバシーへの懸念など倫理的問

[71] ヒアリング資料（抜粋）は <https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。
[72] IoT 推進コンソーシアム　カメラ画像利活用 SWG <http://www.iotac.jp/wg/data/camera/> 参照。

題がクローズアップされるようになったことも一つの背景として、2019 年（令和元年） 7 月に全員が社外のメンバー（役員・従業員はオブザーバー参加）で構成される Ethical Approach to AI （EAA）を設置。

ウ． EAA 発足の経緯

同社内グローバルメンバーから、倫理的観点での事業検証は、グローバルでは非常に重要視されるので検証組織の設立が必要の提案があったことがきっかけ。また、法律上の論点と関連する倫理の問題という法務の観点からのニーズもあった。委員選定にあたっては、国籍やジェンダーのバランスに考慮するというグローバルメンバーからの提案も反映させた。

エ． EAA の位置づけ

具体的事案に即した議論を行う諮問機関としての位置づけ（EAA 総体としての意思決定を行うものではない。）。

（2） 議論

【他のベンチャーと情報共有する取組について】

Q. ベンチャー企業の取組として非常に素晴らしい。こうした取組を他のベンチャーと共有する取組は考えているか。

A. EAA は立ち上げたばかりということもあり、取組を他社に紹介することには至っていない。

【EAA の意義についての自己評価】

Q. EAA のような倫理委員会を設けることは、ベンチャー企業として体力的にも限られているなかで難しいと思う点があるか。

A. ベンチャー企業として体力的には難しい部分もあるが、倫理委員会を立ち上げていること自体がプラスに働いているというのは、役員の共通認識であり、重要度は高いと認識している。

３． 富士通（株）（「AI 開発者における AI ガバナンス」）

（1） ヒアリング概要 [73]

ア． AI 倫理指針の策定及びその概要

2019 年（平成 31 年）3 月に「富士通グループ AI コミットメント」という名称で発表。「人を中心とした AI」、「客観的な AI 倫理指針」及び「AI 倫理外部委員会設置」を 3 つの

[73] ヒアリング資料（抜粋）は <https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

特徴とし、5 項目のコミットメントとして宣言。

イ.　「客観的な AI 倫理指針」について

欧州 AI4People との連携により、科学技術倫理の専門家と連携し、AI 倫理の要件を過不足なく把握。AI4People の AI 倫理 5 原則を踏まえ、同社のステークホルダを含む社会全体へのメッセージに変換して、AI コミットメントを策定。

ウ.　AI 倫理外部委員会の設置

AI 倫理への取組においては自律的な客観性の確保が重要と考え、その観点から評価を受けるため、AI やそれ以外の分野における社外の専門家からなる委員会を設置。なお、現時点で第三者独立機関による監査の仕組みつくりに着手することについては、①AI 技術自体が黎明期であり、方向性が定まっておらず陳腐化も早いこと、②AI 倫理に関する標準化、品質保証などの議論は開始されたばかりで、審査項目も固まっていないこと、③AI は倫理とは関係のない分野での実装も期待されるが、監査先行では全体を萎縮させてしまい、時期尚早と思われることから、まずは自律的な客観性を担保するための外部委員会を設置。

エ.　AI 倫理外部委員会の特徴

AI 倫理をコーポレート・ガバナンス体制に組み込むという仕組みを有する。具体的には、AI 倫理外部委員会は取締役会と情報共有する形態。取締役会は、業務執行の決定や業務執行取締役の監督を行う機関であることから、経営そのものと AI 倫理を結びつけたということが大きな特徴。

オ.　AI ネットワーク社会における AI ガバナンスに関する議論の方向性について

- 議論の対象は、「AI に限らず、ICT 全般の利活用によって社会に発生しうる不都合」と広く捉え、他の ICT に対する規制とのバランスを考慮し、AI に偏らないことが現実的。
- 議論はリスク抑制に偏りがちであるが、今後も激しく進化する AI 技術について、長期的・普遍的に有効なリスク抑制策を現時点で想定することは容易でなく、現時点で懸念されるリスクに対する現実的な対策をしつつも、むしろ、研究やサービスを過度に萎縮させるルール作りは避けるべき。

（２） 議論

【外部委員会の開催以外の活動について】

Q. 外部委員会の開催以外の活動はあるか。

A. 教育活動やお客様から問い合わせがあった際の現場向けのマニュアル整備等、従業員に浸透させる仕組みの構築を進めている。

【AI コミットメントの実効性の確保に関する仕組み】

Q.「AI コミットメント」について、AI の開発やサービス、製品の供給のどの段階において、そのコミットメントの実効性の確保に関する仕組みを講じているか。

A.「AI コミットメント」は、主な商流である B to B to C においてお客様となる真ん中の B と C の方に安心・安全な AI をお届けするというコンセプトであり、AI の開発時点からその理念を念頭におき、最前線のお客様まで届くようにという思いで活動している。

４． 日本 IBM（「AI に関する IBM の取組みについて」）

（１） ヒアリング概要 [74]

ア． 原則及び倫理に関するガイドライン

2018 年（平成 30 年）5 月に「IBM's Principles for Trust and Transparency」を発表。信頼と透明性に関する同社の原則をとりまとめ。

説明性もしくは透明性を確保する取組は同社 1 社が行うものではなく、同社の技術を使う企業もそれらの取組を行っていただく必要があるという強いメッセージ。

そのためのガイドラインという形で、2018 年（平成 30 年）9 月に「AI 倫理のための実践的ガイド」を発表。2018 年（平成 30 年）頃から特に米国で AI の判断に関して一定のバイアス、偏りが発生することによって、社会に不利益を生じさせるといった事例が出てき始めたことを踏まえ、説明性や透明性だけではなくて、偏りのない AI の判断をどのように支援していくのかに関してもフォーカス。

イ． 原則及び倫理に関するガイドラインを踏まえた技術製品提供の考え方

ソフトウェア製品のラインナップに、前述の原則の要素を加える形で製品を提供。Watson API と Watson Open Scale がその要素を非常に色濃く反映。

ウ． Watson API の概要

(i) 概要

音声認識や画像認識、自然言語の理解といった様々な AI のユースケースに対応できる機能群を API で呼び出すという形で部品として提供。その部品群は幾つか学習のパターンがあり、①同社が学習した上で提供するもの、②一部お客様のデータで学習しカスタマイズするもの、③お客様が学習して使用するもの、の 3 つのパターンで API を提供。

(ii) 主な特徴

- (i)②、③の場合にお客様が投入したビジネス固有の情報を同社がどのように扱うのか、

[74] ヒアリング資料（抜粋）は <https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

もしくはそれが他の企業に展開されてしまうのかに関する懸念について、学習成果をどのように活用するかをガイドラインに仕組みとして明確化。お客様が Watson API の学習のために投入したデータをベースモデルの学習に活用するために提供するかどうかについてお客様に選択権を提供。

- 学習データ、もしくは、学習した後の Watson API を使って得られる洞察、もしくは学習データの帰属について、サービス利用開始時の契約で明確化。
- クラウドで Watson API を使用し学習データを投入することに懸念があるお客様に向けて、オンプレミスで Watson API を使用する利用形態も提供。

エ. Watson Open Scale によるお客様開発 AI モデルの信頼性、透明性の確保

- お客様独自に作った AI モデルがどういった仕組みで動いているのか（例えば、精度、偏り、パフォーマンス）のモニタリングを支えるために 2018 年（平成 30 年）にリリース。AI の公平性、説明可能性を、このツールを活用して、AI モデルをモニタリングすることによって担保するという仕組み。
- Open Scale は、競合他社もしくはそれ以外のオープンソースのライブラリを使って開発してもらうような AI モデルに関してもモニタリングする機能、仕組みを提供。企業の中で様々活用いただいている AI モデルを一元的にモニタリングする仕組みとして活用いただくといった考え方で提供。

オ. 説明性・透明性の確保に関して

非常に多くのお客様が関心を示されており、AI の原則をきちんと企業の中で確保しつつ、それを実践していくことに対しての機運が高まっている印象。とはいえ、この Open Scale を使って、AI のモニタリングを実践している企業は、少なくとも日本の市場においてはまだ少数。お客様の現在の関心事は、企業の中で AI を活用するというのがどういうことかについてであり、それがある程度実現でき、その上で AI のモニタリングを行うことができてから説明性や信頼性の確保を進めたいというようなコメントをいただくことが多いのが現状。

ただ、実際に問い合わせが多いというのも厳然たる事実であり、将来に向けて考えた時に、こういった仕組みを企業の皆様に持っていただくことが必要。

カ. 要望

- 企業が倫理や説明性・透明性に取り組むためにはもう少し時間がかかると思うが、説明性・透明性を確保しなければならないというタイミングになってから、それに対してどうやっていくかを考えるのでは、後追いの対応になってしまうとの感触。そういった意味では、今 AI の開発を実践されているお客様に、次のステップとして必要となることを繰り返しインプットはしているが、政府側の動きとしても、AI の活用とその先にあ

る説明性・透明性というところに関して、積極的なメッセージ発信を要望。

- そのためには、説明性・透明性だけではなく、企業としての実際の利活用を促進することが重要。民間の AI 活用のベストプラクティスを収集、公開するとか、日本全体として見た時に、AI を推進していくための機運とそれに対する促進のための後押しを要望。
- データの権利関係の明確化に関しても、日本社会全体としての取組への働きかけを要望。

（２）議論

【フェアネスの考え方について】

Q. データバイアスやアルゴリズムバイアスに関しての話があったが、フェアネスの定義について、スターティングポイントが皆平等である、公平であるという意味でのフェアネスと、終着点を平等にしたいというフェアネスがあり、これらをどのように適用することを考えているのか。

A. 例えば、Open Scale という技術で考えるフェアネスは、あくまでもお客様企業の方々が実際に作った AI モデルの中でどの項目を重視してモニタリングするのかを、お客様に判断いただき、その上で実際にあくまでもツールとして挙げられるものはアラートまでなので、アラートを挙げたものを実際にバイアスであると認知するのか、もしくは、これは然るべき偏りであって、無視して良いものなのかというところの判断に関しては、お客様企業側の運営に任せていくというところを基本的なスタンスとしている。また、フェアネス、公平性に関しては、データとしての偏り、インプットのフェアネスなのか、アウトプットのフェアネスなのかのみならず、社会通念や、環境的な背景にも関わってくるので、基本的には、こういった仕組みを提供させていただく中で、企業独自の公平性を設定していただき、それを、ツールを使って見える化し、モニタリングしていくといったところを支援している。

【透明性の原則：入出力検証可能性（他社の AI モデルのモニタリング）について】

Q. お客様のデータによって、あるいは、お客様の好みによっては、適当でないシステムが作られてしまう可能性もしばしば報告されている。これに対して、企業側として、お客様の責任と言い切るのか、それとも、企業の製品における倫理的な責任をとるかというポジションが、製品の値段にも反映してくると考える。Open Scale では、お客様が制作したモデルに対してマネジメントを行うことと思うが、自社の製品であれば、Ethics by Design のような考え方も適用できると思うが、他社の製品については中身が大体ブラックボックス化しているので、分からないと思う。そこでどのようにしてマネジメントを実現するのか。

A. Open Scale は本番でどう動いているかを見るというところにフォーカスしており、他社の AI モデルのモニタリングについては入力に対してどういう出力が出るかというの

を大量に収集することによって、その傾向を見て、何かしらの偏りがないかをチェックしている。

C. こうした同社の取組は、「AI 開発ガイドライン」の「透明性の原則」における、「入出力検証可能性、説明可能性」の実践に該当するものではないか。

【ガイドラインの浸透について】

Q. お客様が同社に話を聞かれる時に、同社のガイドラインあるいは「AI 開発ガイドライン」を認知しているか。

A. 当社が出しているガイドラインについて、お客様が認知していることはまずなく、また、政府が出しているガイドラインに関しても、直接その言葉を引用する形でお話をいただくということは、なかなかないのが実際であるが、お客様から漠然と世間一般で trust または explainability に対するリスク、課題意識が高まってきているという発言があるため、政府から出されているものなどを勉強しているのかなと感じるところがある。

【Explainable と性能とのトレードオフについて】

Q. Explainable AI が流行ってきているが、explainable にするほど性能が落ちてくるというトレードオフの関係にあるという認識だが、そこのところはどのように対応されているのか。

A. 説明性を確保するために、AI の作り方をホワイトボックスにしましょうというような、いわゆるホワイトボックス AI による explainability の担保ということをやっている企業も幾つか見受けられる。ただ、ディープラーニングが得意とする画像や音声といった大量のデータを分析して、分析の AI モデルを作るというところに関して、なかなか explainability をホワイトボックスで担保するのは難しい。当社の技術は、モデルの中身に手を入れるのではなく、Open Scale でやっているような外側から見るといったアプローチで、AI モデルの作り方自体に制約をかけることなく、外側から説明性を担保する方法、すなわち、ディープラーニングの技術をしっかりと活用しながら、説明性をお客様に享受いただくというアプローチで支援することを中心に進めている。

【公平性や信頼性に対するコミットメントと制度的対応について】

Q. 公平性や信頼性はお客様が設定をされて、それを検証するという形であるということだが、企業の社会的責任性を考えた時に、もし、お客様企業が何か問題を起こせば、それはそのシステムを提供した同社も同じく名指しされるとは思われるが、その時に、社会的責任に関して、あくまでツールだと言いつつも、御社として、コミットメントをしているのか、あるいは、する用意があるのか。会社の立ち位置を教えてほしい。また、その時に開発の担当者だけでは公平性や信頼性に関して分からない可能性が高くて、最

近では、専門の人を開発に混ぜるとか、あるいは責任者を 1 回通すという対応があるが、制度としては作られていないのか。

A. 実際に、カスタマイズした AI モデルをお客様と一緒に作っていく中で、AI の振る舞いに関してモデルを作ったから、そのまま一気に公開ということはやっておらず、プロジェクトでテストを何度も何度も繰り返しながらやっていくという形をとることで、お客様と一緒になって AI の挙動に関して不適当な動きをしないようにというところを担保するアプローチをとっている。また、現状として何かの特定の試験を必ずパスさせないと公開させないとか、そういった取組にはまだ至っていない。

５． （株）NTT データ（「NTT データグループ AI ガバナンスに関する取り組みについて」）

（１） ヒアリング概要 [75]

ア． AI 指針策定にあたっての背景

AI 原則に関する様々な活動が広まってきているなかで、2019 年（令和元年）6 月に G20 茨城つくば貿易・デジタル経済大臣会合において、日本の「AI に関する社会原則」の公表が予定されることとなり、同社としてもグローバルに AI を広く展開する IT Service Provider として、AI 指針の提示が責務と考え、2019 年（令和元年）5 月に発表した。

イ． 同社グループにおける AI ガバナンスに向けた取組

現在発表している「AI 指針」は AI 開発における理念の部分であり、内容的には一般的なものとなる。

今後は、「AI 指針」を踏まえた、より具体的な AI 開発のプロセスを整備する予定である。AI のシステムを開発するにあたり、お客様と一緒にどういう開発のプロセスで AI システムを作っていくのかを最初に意識合わせをすることから、開発のプロセスの中でどのような内容を具体的に確認していくのかというところのブレイクダウンをしていかなければいけない。そのため、今年度中に AI サービス実現のプロセスとチェック観点を整備した「AI 開発プロセス」を作っていこうとしている。ただ、プロセスだけでも、実際にどうしたらいいかというのはなかなか分かりにくいので、さらに具体的な課題解決方法、確認方法を整備した「AI 診断手法」を一般的なツールの利用や不足するツールの自社開発も含めて検討している。

ウ． AI 指針策定のプロセス

同社のグループビジョン「Trusted Global Innovator」を実現するためにどのような指針であるべきなのかということを最初に検討し、その上で、国や他社が既に発表している原則

[75] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

や指針を参考にしつつ､その上で国内外のグループ内の意見を入れてAI指針を完成させた。

また、AI 指針の発表にあたっては、取締役会に付議し、承認を受ける形をとり、経営層まで含めてコンセンサスを取った上で発表している。また、今後、AI 指針は社会の変化や技術の進展に合わせてブラッシュアップしていく予定である。

エ. 事業活動での取組

一例であるが、海外の AI の専門家と連携して、「AI Center of Excellence」を発足。コアメンバーは、スペインが 90 名、日本が 50 名で計 140 人。メンバーは 7 カ国 8 社が参加している。

現在「AI Center of Excellence」を中心に、顧客向けの AI 倫理に関するアセスメントや方法論を体系化した「AI Ethics Framework」を整備している。これはガバナンス、アルゴリズム、オーガニゼーション、ソサイエティそれぞれの観点での確認事項をフレームワークとして整理していくものである。

オ. 「AI 開発プロセス」と観点の例

AI 開発は、従来の大規模システムのウォーターフォール型開発のように進むことはあまりなく、アジャイル型に進んでいくことが多い。「AI 開発プロセス」はアジャイル型の AI 開発におけるフェーズごとの具体的な確認項目と確認方法を定義している。

例えば、「AI 開発プロセス」におけるデータ収集では、データ量やデータの質などの重要確認項目を整理したアセスメントシートと、AI バイアス可視化・検知ツールのようなチェックツールの組み合わせにより、開発者に具体的な開発プロセスを提示している。「AI 開発プロセス」を利用することにより、実際にお客様と一緒に作っている AI のサービスが AI の原則に本当に則っているのか、お客様の要望を実現できるのかということを確認することができると考えている。

（２） 議論

【指針等の企業活動での活用における制度的対応について】

Q. 指針やプロセスを実際に企業活動で運用する上で、専門的な方に開発に入ってもらうとか、あるいは開発があった時に、具体的にどこかのセクションで確認するとか、そのような制度的対応についてどのように考えているか。

A. 組織やそういう仕組みでチェックする、確認するということは現時点では考えていない。開発プロセスが社内標準の Methodology になっているためである。その上で、どれだけ適用するかはお客様との話し合いの中で決めていくが、その中で困ったことがあれば、技術部隊や、品質保証の部隊もおり、そこに相談に来るというのが通常の開発でも起こっているので、そういう形で進めていくことを考えている。

【フェアネスの基準について】

Q. フェアネスについてどのように基準を作っているのか。

A. お客様との話の中で、これで良いのかということを確認していくというのが基本的な考え方で、サービスとして提供する時に、いきなり大規模に行うことは通常なく、少しずつ提供しつつ反応を見ながら、リスクを下げていくというのが現実的な解であり、公平性担保に関する基準は現状ない。

【Data minimization について】

Q. お客さんから情報を集めなければならない場合でも、不必要に多くのデータを集めない、Data minimization という考え方が最近非常に重視されていると思うので、この点について、同社の理念の部分でどう考えるのか。また、顧客によって集めるべき情報が違うと思うが、どういう考えで進められるのか。

A. Data minimization という考え方を理念のところに入れるべきかどうかは今後の検討項目だと考えている。現実に、お客様とのやり取りの中でそういった議論も起こっており、お客様の方も出したくないといっているし、私たちも不必要なデータはいただきたくない。なお、データをいただく際には、データを加工する、一部マスキングするというツールを提供しており、特に機密性の高い医療系、ヘルスケア系のデータではツールによるマスキングがデータ提供の前提となっている。

【データガバナンスに投下するコストの実態について】

Q. 各ベンダーがデータガバナンスに投下しているコストの実態は把握していないが、今後全体のコストに占めるデータガバナンスに投下するコストについては IT ベンダーにとって結構深刻な問題になっていくと考えられる。その点についてどのように感じているか。

A. データの種類や業界によって投下コストは異なるという感覚を持っている。機密性の高いデータを扱うことが多い医療、ヘルスケア、ライフサイエンスの分野はデータガバナンスに対して、非常にセンシティブなことになっているのは間違いない。ただ、現時点においては具体的な割合は把握できない。また、同社はデータ管理プラットフォームを提供しており、データ管理のガイドライン・チェックリストは適用しているが、医療、ヘルスケア、ライフサイエンスの分野では、お客様の方が厳しいデータ管理のガイドラインを持っている場合があり、その際には、お客様のガイドラインに従ってプラットフォームとサービスを作っている。また、リテール系も最近はかなりデータの取り扱いについてセンシティブだというのは間違いない。

【データガバナンスに関する情報共有】

C. データガバナンスに投下するコストの実態について業界で共有できるような場ができ

てくると、比較的健全にいくのではないか、あまりにもケアをしていないベンダーが出てくると、そこが産業的にビハインドする可能性があり、同社は、我が国として最大のシステムインテグレーターであり、そういう役割も期待したい。

A. こういう時にこのようにやっています、というベストプラクティスを収集し、共有するというのは良いやり方だと考えている。AI の原則はたくさん出ているが、これであれば絶対に大丈夫だということは言いにくいので、こういう条件のこういう制約がある時に、これをやってうまくいっていますという事例をある程度集めないといけないのではないかという感覚を持っている。お客様への配慮が必要な部分ではあるがある程度抽象化した形で事例として提供するということはやっていきたいと思っている。また、「AI 開発プロセス」や、「AI 診断手法」というのは、幾つか事例をやってみて、それをある程度汎用化して作るというパターンになってくるので、こういったものが世の中にある程度認めてもらえれば、その中にデータのガバナンスの話も入っているという形になると、一番嬉しい。さらに、例えば AI の QA（品質保証: Quality Assurance）で QA 4 AI という活動があるが、素晴らしい考え方が出されてきており、そういったところもどんどん参考に入れて、そして事例ベースでどうやったという議論をしていくべきではないか。

６． 沖電気工業（株）（「AI 実用化に向けた環境整備～「OKI グループ AI 原則」の制定～」）

（１） ヒアリングの概要 [76]

ア． AI の取組

2010 年代くらいまでは研究所の中で AI を開発し、製品とは少し遠い話。2010 年代に入り、AI が事業中心の方にフェーズが移行。製品としては 2000 年代から画像認識などでは機械学習を使った製品は販売していたが一部であり、それほど大きな売り上げはなかった。最近は事業部門の大きな主力製品として AI を活用していくという流れ。

イ． AI 環境整備プロジェクト

徐々に社内でも AI の事業・商品が出つつあるなかで、製品をお客様へ届けるときの、契約あるいは品質保証は個別の部門がそれぞれ最適と考えて対応してきているが、横通しで見ると、それぞれに温度差があり、事業部門自体もこれでいいのかという不安もあった。このため、社内の統一的なルールが必要ということで、2019 年度（令和元年度）に AI 環境整備プロジェクトを全社で立ち上げ。

AI を活用した事業推進には、技術面と非技術面があるが、この環境整備プロジェクトは主に非技術的側面、倫理・品質保証・契約などにフォーカスしたプロジェクト。

---

[76] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

プロジェクトには、AI のビジネスに関わるほぼすべての部門が参加。AI に関する施策を運用する、作っていく側と、それを使う側が一緒になってガイドラインを作ることによって、実際に無理なく運用できる、現実的な仕組みをここで設計することが目標。

ウ. AI 原則策定のプロセス

既に出されている各国・各団体の採用項目をすべて洗いだし、これと同社の企業行動憲章・行動指針等を突き合わせ、同社グループが共有する価値観との整合性を保ちつつ特長を表現する形で 2019 年（令和元年）9 月に制定。

エ. 原則の位置づけと制定の狙い

まず一つは社会的責任を果たす為の基盤であるということと、すべての AI に関する活動をこの AI 原則に基づいたものとする為の原則であるということを位置づけ。

特定の狙いとして、基本的な考え方を提示するということと、同社が共有する価値観として、すでにある行動指針等を更に補強するということを明確化。

オ. 今後の取組

AI 原則の策定を受けて今後はこれを遵守する体制をどのように作っていくか、どのように遵守させていくかというガイドラインを作っていく。

今後の運用に向けて、事業推進の中で様々な AI の製品を出すにあたり、押さえておくべきリスクなどをどう管理していくかを考えていく。昨年度については試行運用ということで、AI の商品を多く出している情報通信の事業部門に対して、お客様に製品を届けるにあたり契約はどうか、品質はどうか等様々なチェックプロセスを実施するプロジェクトにおいて、これは AI 関連案件かどうか、その中でどういったことに配慮しないといけないかということの事前チェックを実施。この結果を踏まえ、今年度中に本格的な運用を行う予定。

試行では、現場に与える負荷がどれくらいあるかと、コストがどれくらいか等を確認する。本推進会議「AI 利活用ガイドライン」のトレードオフの例[77]にもあるように、現場の人は余計なプロセスが入るとやはりビジネスを阻害されるようなイメージを持ち、なぜやらなければいけないかということを必ず言われる。できるだけ現場に対して負荷を与えないような仕組みを整える一方でお客様に後になって実はこうでしたとならないように、必要な対処を行う等、トレードオフを洗い出すのが 2019 年度（令和元年度）の試行の目的。これによって AI が面倒な商品である、コストがかかるといわれることも阻害要因となるので、この試行を通して議論した上で、2020 年度（令和 2 年度）からは無理のない運用ができるよう体制とルールを構築する。

---

[77] 本章１．参照

（２） 議論

【開発原則や利活用原則の認知度、FAT[78]への関心について】

Q. 今までは演繹的にルール・原則を作ってきたが、実装段階になると恐らく帰納法的に今度はプラクティスを集めてお客様の声を聞きながら、例えばフェアネスなんかも聞きながら一緒に作っていって、少しずつ展開していくと、恐らくこの方向性にあるのかなと思う。そこで同社は B to B の業務がメインということだが、お客様である企業がこの政府の開発原則や利活用原則を知っているかどうか、いわゆる FAT についての関心がお客様企業も高いのかどうか。

A. お客様によると考えられ、金融関係のお客様はやはりフェアネスは非常に気にすると思う。お客様全般に見るとメッセージとしてこちらから伝えないといけないと考えている。

【「社会的受容可能性」と社内での浸透について】

Q. 原則策定の目的の 1 つである「多くのお客様や社会に受け入れていただける AI 商品などを提供する」という「社会的受容可能性」を言うことによって現場でも受け入れられやすくなる、というようなことがあるのか。

A. 現場も人によるということで、先を読める SE や営業の方はそういったことは関心があって、理解してもらえるということがあるが、なかなか理解が難しい人はそういった社会情勢を言ってもそれだけでは受け入れられない。その辺のリテラシーの教育はこれから地道にやっていく必要がある。

７． Microsoft（「ビジネスと責任、AI を取り巻く課題と取り組み～倫理と AI の可能性～」）

（１） ヒアリングの概要 [79]

ア． 協働による課題解決

AI 活用は歴史が浅く、未知の課題が多く残されており、安全・安心な AI 活用の為には様々な領域の組織・団体との協働が不可欠。

イ． Microsoft AI 原則と AI 倫理審査ボード

2017 年（平成 29 年）に AI に関する原則をいち早く策定。現在はこの原則に基づいて AI 倫理審査ボードというものを設けて企業活動を統制。透明性と責任を 2 つのベースとし、透明性の上に①公平性、②信頼性と③安全性、責任の上に④プライバシーと⑤セキュリティ、

---

[78] Fairness（フェアネス）、Accountability（アカウンタビリティ）、Transparency（トランスペアレンシー）の略。

[79] ヒアリング資料（抜粋）は <https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

⑥インクルーシブという 6 つの原則。製品の研究開発や実際のサービス提供において、いかに 6 つの原則に沿ったビジネスを行っていくのかということが基盤。

ウ. 倫理と責任ある AI（公平性と信頼性）

2018 年（平成 30 年）に NY タイムズで紹介された、人種によって年齢設定の品質が異なるという非常に有名な例。データの偏り、データのバイアスといってもよいが、それによって AI を活用することに関する公平性、信頼性への疑念が生じた大きな象徴的な事例。

公平性についてはシーン・状況において様々に異なる。同時に AI のアルゴリズムだけではなくて学習に使用するデータの分布、偏り等についても目的・用途に応じて非常に慎重に精査しなければならないということを明らかにした事例。AI ベンダー・IT ベンダーは、アルゴリズムやチューニング等に注力しがちだが、実サービスとして提供する際にはこの対象とそれに基づいた公平性とはいったい何なのかということを踏まえたデータを含む精査が必要。

エ. 倫理と責任ある AI（透明性）

データの収集、各種 AI システムの構築に至るまでどういう考え方に基づいて作成を行ったかということについては、この考え方を明らかにすることが非常に重要。（肺炎による死亡リスクの予測の研究において喘息患者の肺炎による死亡率が非常に低い事例を紹介しつつ）実際の目に見えるわかりやすいデータだけでは本質が見えないケースも非常に多い。ビジネスという観点から見た場合、AI を用いたシステムを開発する上では、対象とそのデータの意味・本質をしっかり理解するということが重要。

オ. 倫理と責任ある AI（プライバシーとセキュリティ）

（米国のスーパーマーケットターゲットで女性のネットでの検索などの行動履歴からその女性の父親よりもお店の方がその女性の妊娠を早く認識していたという事例を紹介し）データと分析、モデル化等から実際には開示していない、内に秘めている個人情報も外部から明らかになる可能性があるということを想定した上で AI システムをどのように運用していくかが 1 つの課題。

カ. 倫理と責任ある AI（説明責任）

顔認証に限らず、AI を活用していくことによってこれまで不可能であった極めて膨大なデータを用いた分析・モデル化といった様々なことが可能。望ましくない用途に AI が使われる事によって様々な社会的影響が出る領域も考えられる。同社では研究開発を進めるとともに、慎重に評価・検討するという方針。

キ. AI システムの責任ある開発と使用における考慮事項（倫理審査ボードの具体的評

価）について

- システムの目的。同社のビジョンとしても、製品サービスが人々の可能性を最大化するということを掲示。当然 AI の研究開発や提供が人々や社会に良い影響を与えるものなのかどうかということを第一の評価ポイント。
- テクノロジーレディネス。倫理等とは少し異なるが、何でも AI を使うというのではなく、技術的に AI を使うことが正しいのかどうか、適切なのかどうかという判断・評価。
- 品質と信頼性。特に品質の維持が 1 つの大きなポイント。AI のモデルを構築するにあたり、集めたデータが非常に多岐にわたり、しかもボリュームが非常に大きいことで、システム開発のデータ収集に膨大なコストがかかる。品質と信頼性を維持するということにおいて、AI を使う場合、対象によってはコストがかかり、それを維持することが困難になるケースも容易に考えられる。やみくもに AI を使うというのではなく、持続性を慎重に評価することも重要。
- 慎重な AI 活用。AI を用いたシステムは用途に応じて慎重な評価と対応が必要。AI を用いたシステムを開発しサービスとして提供する際には対象は何か、人や社会への影響はどうなるのかということについても事前に慎重に評価。
- AI によるサービスの拒否。それが起きうる用途なのかどうなのか、差別・格差を助長するような結果をもたらさないか、そして個人の自由とプライバシーを侵害することはないのかといった観点から評価。

ク． AI とデータを取り巻く課題（AI データ活用コンソーシアムの取組）

円滑なデータの流通の実現を通して社会課題の解決を促進することが重要（例としての障がいのある方に向けた AI データ活用、高齢者に向けた AI データの活用等の事例については、**第５章　消費者的利用者に関する取組**で記述）。

細分化されたデータ提供者、異なるライセンスの考え方、個人情報、計算リソースとの連携など、AI 研究、オープンイノベーション、ソリューション化（商用化）には解決すべき多くの課題が存在。

ケ． データ流通基盤に求められるさまざまな要素と課題

(i) データが持つ異なる性質

データが広く流通活用され、様々なベンダーが様々なソリューションを提供して社会課題の解決に様々な団体が貢献できるということが理想。しかしデータを見た場合、従来のデータ分析・活用とは異なる性質。機械学習の場合では学習した後のモデルにデータが形を変えてある意味永続。データの持つ価値がこれまでのデータ流通と全く異なり、技術特許に近い特性を持つ。通常技術特許の場合は、技術を適用する製品の商流をヒアリングして交渉の末、契約内容と価格を決定。一方、AI のモデルについては、学習したデータそのものを用いて作成・構築した AI システムがサービスとして提供もしくはデバイスに組み込まれて販

売されるなどするため、最終的にデータの持つ価値は商流が確定しないことには確定しない。いわばデータそのものが 2 次的・3 次的な価値を持つ。最近の傾向として、AI リテラシーの高いデータホルダー・データ取得者ほど、技術特許と同様の認識。これが反対にデータ流通、契約を含めて難しくしている状況。

データ流通を促進する場合、データ取得者と AI 開発者、でき上がったものを使うユーザーそれぞれが異なるケースも増えていくと想定。そういったことを想定した契約モデルが存在しないことから、AI データ活用コンソーシアムにおいて契約モデル、テンプレートなどについて議論・検討し作成するという作業を実施。

(ii) ソフトウェアプログラムと製造物責任

ソフトウェア開発におけるアルゴリズム実装の場合は、欠陥があった場合の責任所在の明確化が比較的可能。

他方、AI を用いたシステムの場合、AI システムの適用対象によっては非常に重要な用途。アルゴリズムそのものに問題がある、改善の余地があるということもあるが、説明可能な AI を技術的に確立するにはまだまだ道のりは長い。

(iii) 現行 AI の品質におけるデータプロベナンス（各種データそのものに対する責任の所在）

画像を学習データとして使用する場合は必ずアノテーションを使用。ベンダーは多くの場合アノテーションをアウトソース。例えばそのアノテーションベンダー事業の中に悪意を持った作業者がいたと仮定し、実際に AI のシステムが問題を抱えていた場合、データまでさかのぼって悪意のあるデータが混入していたのかどうか、AI の場合には学習に用いたデータ、または非常にわかりやすいアノテーションのラベリング作業において悪意を持って作業した人間がそこにいたのかどうか、いた場合には誰だったのかということをトレースできるのかということが今後重要。

大きくレベル 1・2・3 と分け、例えばレベル 3（身体などに様々な回復不能なインパクトを与えるような AI システムの用途）の場合、学習に用いる元のデータやラベルデータの誰がラベリングしたのかなどの出どころをさかのぼってしっかり説明責任を果たせるということも今後は重要。

また、品質に関して、データの環境の変化がデータの陳腐化、新しいデータが必要になる、そして AI の品質にも影響が出るということにつながる。これが環境の変化によるものなのか、何かそこに悪意が入っているものなのかということも含めて慎重に評価・検討することが必要。

（２） 議論

【悪意のあるアノテーションと製造物責任】

Q. 学習データのアノテーションに悪意が入っている場合、あるいは悪意でなくてもバイアスが入ってしまう場合はしばしばあると考えられ、これは現代の AI にとって非常に大きな問題。そのような場合、製造物責任というのをどこまで問えるのか。

A. 悪意のあるアノテーションと製造物責任に関しては、これは対象や内容によっては責任が生じると考えている、製造やシステムを開発する側としては学習データをそれなりに精査して誤ったデータが混入していないことを確認する義務があると考える。当然それなりに誤ったデータが混入していないことの精査、品質を確認する努力をしていたにもかかわらず問題が生じてしまった場合と全くそれをせずに問題が生じてしまった場合とでは大きな違いがある。そういった意味で AI データ活用コンソーシアムではアノテーションの作業をどこの誰がやったのか、データプロビナンスやトレーサビリティというものをしっかり保証や担保ができる仕組みづくりということも検討している。

【AI によるミスラベリングのチェック】

Q. 非常に大量のデータを学習データとして使うなかで、それを人間の目で見るというのは二度手間になってしまう上にとても大変。むしろ AI のようなテクノロジーを使うのであれば、悪意のあるデータの候補をチェックするような AI テクノロジーの導入の仕方を考えるべきではないか、反対にそれをやらなければ製造物責任をしっかり果たしたと抗弁できなくなるのではないか。

A. 取組例として、膨大なデータを学習させる際に、膨大な学習データをすべて目視でチェックするということが現実的だとは考えていない。一方で限定的にチェックをした、ミスラベリングを含まないデータをある程度集めるということも可能。したがってそういったものを用いて学習させて、それ以外のラベルデータをある程度用いて品質チェックするという再帰的にまわしていくというような手法について現在リサーチ中。

【AI のデータの透明性】

Q. 例えば保険等の個人データの透明性について、間違って発表されている可能性も出てくると思うが、どのように担保されているのか。

A. 誤ったスコアリングによる判定で不利益を生じるような結果をもたらすような目的でシステムは作らない。そこはリスクがあるためそういった分野について、AI を用いたシステム開発をしないという判断。説明責任を果たす、透明性を確保するのが難しい分野のシステム開発やサービス提供はしないという判断。

【製造物責任の検討の経緯】

Q. 製造物責任の検討の背景は。

A. 悪意ある学習データが気づかれないうちに少しずつ混入されていくことでモデルがど

んどん変質していくというようなリスクも 3～4 年前から対処しており、今日 AI の様々なセキュリティ上の脅威と言われているようなものについても実際リサーチ分野ではかなり以前から検討されている分野。それと合わせて Tay という現実的な問題が発生し、緊急でそれを専門に作業・リサーチするチームを立ち上げて取り組んできたというのが経緯。

【アノテーションのモニタリングについて】

C. 数年前にシリコンバレーの企業に行った際に、アノテーションあるいはデータクレンジングのモニタリングがしっかりできているのか見ていて疑問で、今後極めて重要になると思う。クラウドソーシングの契約、特にブローカーを担当するような企業がどういう責任や知財の管理をするのか今後 AI が普及すればするほど重要になるとのコメントがあった。

A. ラベリングやアノテーションのベンダーの調査をしたところ、品質保証を取り下げることを含めてそこについて明言しているアノテーションベンダーは皆無。その検査・審査・チェックをするというサービス・オプションとして提供しているところはあったが、その結果については保障しないというところが基本的にすべてであった。

８． 匿名（「民間企業有志による AI 利活用のための支援ツールに関する報告」）

（１） ヒアリング概要 [80]

ア． 趣旨

民間企業有志による研究活動（以下「本研究活動」という。）として、AI 利活用のための支援ツールとして、AI の導入を検討している企業がセキュアに導入するためのチェックリスト策定について検討。

イ． 経緯

本研究活動では、多くの企業で AI の導入について関心が高く、導入を検討していることを背景とし、「AI 利活用ガイドライン」を参考にするとともに、日本と EU の AI 利活用の取組状況を比較した上で、日本でまだ着手できていない領域である AI を導入する際のチェックリストを研究の対象と設定。

ウ． 成果物としての AI 導入支援ツール

AI サービスプロバイダー及びビジネス利用者を想定ユーザとし、誰でも簡単に AI を導入可能とする観点から、①AI Assessment List（以下「本件 Assessment List」という。）、

[80] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

②本件 Assessment List 利用説明書及び③チェック項目解説書から構成。また、導入者の立場や導入のフェーズによってリストの絞り込みを可能とするなどを工夫。

エ. 検証

本研究活動では、本件 Assessment List を実際の現場（製造業 1 社、保険業 1 社）で使ってもらうことでフィードバックを得ることとし、118 のチェック項目のうち、84%が有効な項目となったと検証。また、重複している質問事項があった等の指摘事項を踏まえた改善プロセスも実施。さらに、EU「Trustworthy AI Assessment List」と本件 Assessment List を対比することにより、本件 Assessment List だけにある項目（理由は EU においては GDPR でカバーされていると考察）がある一方で、EU「Trustworthy AI Assessment List」だけにある項目があることを確認し、当該項目については、本件 Assessment List に追加する改訂を実施。

オ. 今後の展望

AI をビジネス利用していくためには、「AI 利用の指針」に基づいた「自己点検・自己評価」をしていく仕組みは必要であり、さらには、「中立的立場による評価」の仕組みも必要。

「AI 利用の指針」としては「AI 利活用ガイドライン」があり、「自己点検・自己評価」は本件 Assessment List がその役割を果たす。

「中立的な立場による評価」は、監査が考えられるが、監査法人などからは具体的な表明はされていない状況で。クラウドサービスの評価と同じように、今後、AI の利用が広がり、市場の要求に応じて制度が確立されることを願う。

当研究会の活動を継続し、検証結果を踏まえて、本推進会議と連携して AI 利活用の自己点検・自己評価をしていく仕組みの確立を推進。

（２） 議論

【今後の取組への期待】

C. 非常にすばらしい取組である。

C. 本研究活動を進めていただき AI 利活用ガイドラインの改善にもつなげていければよい。

C. 我が国におけるボトムアップの取組として OECD 等に報告できるとよい。

C. EU では Assessment List にかなり力を入れており、Assessment を通じて原則などもリバイスし、データベースをしっかりと作ろうとしており、その意味で非常に貢献してもらえる。

９． とりまとめ

（１） AI 原則等の策定の意義

AI 原則等の策定・活用に意欲的に取り組んでいる企業は、「公平性」、「アカウンタビリティ」、「透明性」など AI の開発等において懸念される点について、明確なメッセージを社会に対して発信している。また、AI 原則等の策定にあたっては、その企業理念等を反映させている特徴もみられる。AI 原則等の策定そのものは直接企業の収益に結びつくものではないものの、AI の開発等に対する企業としての基本的な方針を示すものとなっており、企業のステークホルダにとっては、AI の開発等において懸念される点についての不安を払拭し、AI の開発等の取組に対する信頼を醸成することにつながるものとなる。こうした意義も踏まえ、今後の取組として以下の取組が考えられる。

ア． 企業における AI 原則等策定の取組のフォロー及び周知・PR 等

本章「１．**議論の出発点**」で述べたように AI 原則等の策定の取組が十分な広がりをみせていないなかで、こうした AI 原則等の策定・活用の意義を企業活動に有益な取組という観点から引き続き本推進会議におけるヒアリング等を通じて事例を収集するとともに周知・PR していくことが必要と考えられる。併せて、AI 原則等の策定の参考として資するために、「AI 開発ガイドライン」及び「AI 利活用ガイドライン」の周知を引き続き行うことが重要である。

イ． AI 原則等の策定に関する国際的な議論の動向のフォロー

また、AI 原則等の策定にあたっては、AI に関する取組がグローバルな性格を有することから海外の関係団体と連携するなどの取組は有用である。そのため、国内の動向のみならず、海外及び国際的な議論の動向をフォローすることが重要であるが、こうした動向を時宜に応じて的確にフォローすることは個別の企業によっては必ずしも容易なことではない。そこで、本推進会議として、引き続き海外及び国際的な議論の動向をフォローし情報提供していくことが必要である。また国際的な議論の動向をフォローするためにも OECD をはじめとする国際機関に我が国の取組状況を積極的に発信していくことも必要と考えられる。

（２） AI 原則等の AI 開発・利活用における活用

AI 原則等は企業の AI 開発・利活用に関する理念を表すのみならず、実際の開発・利活用における指針として用いられることで具体的な機能を発揮することができる。例えば、AI の開発において公平性や信頼性についてどのように対応するのか、説明可能性と精度のようなトレードオフの関係にどのように対応するかなどは重要な課題であるが、こうした課題についての判断の前提となるものである。こうした意義を踏まえ、今後の取組として以下の取組が考えられる。

ア. 企業における AI 原則等の具体的な活用事例の収集及び周知・PR

AI 原則等を実際の AI の開発事例等においてどのように活用しているかを知ることは、AI 原則等の策定を検討している企業にとって有益な情報となる。今後、こうした情報を共有し参照できるようにすることは、AI 開発・利活用における事業判断を支援する観点からも重要と考えられる。そのため、本推進会議におけるヒアリング等を通じて引き続き事例を収集するとともに情報発信・情報共有を進めていくことが必要と考えられる。

イ. 具体的診断ツールとしてのチェックリスト等の研究

AI 原則等を実際の AI 開発・利活用に活かしていく上で、AI 原則等を踏まえたチェックリスト等を策定することは、AI 開発・利活用の判断手法の客観性・統一性・検証可能性等を確保する観点から重要と考えられる。こうしたチェックリスト策定の動向についても本ヒアリングにおいて発表があったことは参考となる貴重な取組といえる。他方、EU においても Assessment List を策定、試行し、その試行結果を踏まえ本年夏頃を目途に改訂する取組を進めている。こうした動向を踏まえ、本推進会議では、引き続き EU 等海外の動向のフォロー、国内でのチェックリスト策定の事例の収集を行うとともに、具体的診断ツールとしてのチェックリスト等の研究をしていくことも有益と考えられる[81]。

（3） 安心・安全で信頼性のある AI の開発等に必要なガバナンス体制[82]

AI 原則等の策定にとどまらず、AI 原則等の実施を確保するためにはガバナンス（仕組み）が必要となると考えられる。ガバナンス体制としての自己点検・自己評価の仕組みとして、本ヒアリングでも紹介されたように、外部の多様な人材から構成される社内委員会を設置するなど、工夫した取組が見られる。ガバナンスについては、どのような形で担保されるか、また、どの範囲まで及ぼすものなのかについては様々な形態が考えられるものであることから、今後の取組として以下の取組が考えられる。

---

[81] 例えば、東京大学では、AI サービス提供にあたって、そのリスク要因を技術的要素、サービス提供者の行動規範要素、ユーザー理解・行動・利用環境要素の三層に分類したリスクチェーンモデルを提案している。このモデルを用いてリスクシナリオを作成し、関連する要素の関係性（リスクチェーン）を可視化することによって段階的なリスク低減の検討と効果的・効率的なリスクコントロールを検討し、ステークホルダ間での対話の促進とベストプラクティスの蓄積をするプロジェクトが進められている。
（東京大学未来ビジョン研究センター、AI サービスのリスク低減を検討するリスクチェーンモデルの提案 < https://ifi.u-tokyo.ac.jp/project-news/7079/ >）

[82] AI 原則を実装していくことについて、コーポレート・ガバナンスの議論を踏まえ整理したものとして、総務省情報通信政策研究所情報通信法学研究会 AI 分科会令和 2 年度第 1 回会合における小塚荘一郎構成員（学習院大学法学部教授）発表資料「「AI 開発原則・利活用原則の事業者による実施とコーポレート・ガバナンス」参照。
<https://www.soumu.go.jp/main_content/000689960.pdf>

ア. 自己点検・自己評価の取組事例の収集・PR 等

自己点検・自己評価の取組の普及を図るため、本推進会議として「AI 原則等実施のためのガバナンス」の参考例となる取組事例を収集し、周知していくことが必要である。また、とりわけベンチャー企業においては、社内リソースの優先度等の観点からこうした取組に躊躇することも考えられる。そのため、関係団体等と協力することを通じ、ベンチャー企業との意見交換会を実施していくことが有益と考えられる。

イ. 外部監査についての検討

外部監査については、まだ取組事例が見られていない。引き続き外部監査の在り方について本推進会議において関係者からヒアリングを行い、検討を進めていくことが必要と考えられる[83]。

ウ. ガバナンスの実施内容や課題を共有するための公開された議論の場の設置等

以上のとおり、国内で AI 原則等を策定・活用したりガバナンス体制を構築したりする企業（及び企業グループ）が少しずつ出てきている中で、それぞれの実施内容や課題を共有するため、国内シンポジウムを開催する等公開された議論の場を設置することも有益であると考えられる。また、本推進会議がそれぞれの実施内容や課題を共有するためのプラットフォームの１つとなっていくことが必要と考えられる。

（４） 「AI 利活用ベストプラクティス」の策定

AI 原則等の策定をはじめとする倫理的取組やガバナンスに関する取組の意義について徐々に浸透しているが、ビジネスの現場の感覚として、ビジネス利用者である企業等をはじめ最終利用者において AI を活用することがどういう意義、メリットを有するのかをまず理解してもらうことが必要との指摘があった。こうした AI 利活用の有用性を理解してもらうことは AI の社会実装を進める上での大前提であることから、本推進会議において意欲的に AI の利活用を進め、経営の改善に活かしている方々からヒアリングを行い、AI 利活用の有用性をわかりやすく紹介した「AI 利活用ベストプライティス集」を策定することが必要と考えられる。その際、特定の業種に偏らずに広く多様な業種から事例を収集することが必要と考えられる。

---

[83] 2018 年（平成 30 年）12 月、パーソナルデータ＋α研究会により提示されたプロファイリングに関する提言案（< https://www.shojihomu-portal.jp/nbl1137pc >）がチェックリストや外部監査を検討していく上で参考になるのではないか、また、これをベースに、業界団体（ピープルアナリティクス＆HR テクノロジー協会）が人事（HR テック）領域における AI 利用（プロファイリング）の原則を公表しており、他分野の適用の意味でも参考になるのではないかとの意見があった。

（５） その他

本ヒアリングにおいては、上述のヒアリング概要、議論のとおり、「安心・安全で信頼性のある AI の社会実装」の課題について、AI 倫理を出発点としつつ多岐にわたるものとなっている。こうして提起された課題については、必要に応じ他の関係団体等とも連携しつつ、引き続き検討を行っていくものと考えられる。

第４章 ビジネス利用者における取組

１． 議論の出発点

AI ビジネス利用者は、AI 利活用ガイドラインにおいて「最終利用者のうち業として AI システム又は AI サービスを利用する者」と定義されている。こうした AI ビジネス利用者として『報告書 2019』では、AI サービス（例：医療用 AI クラウドサービス）を利用したサービス（例・医療サービス）の提供を行う医師、AI システム（例：融資審査システム）を利用したサービス（例・融資）の提供を行う金融機関、AI サービス（例：異常検知サービス）を利用するメーカー等が挙げられている。これらはあくまでも例に過ぎず、AI は、これまでの ICT と同様にあらゆる産業活動・社会生活面で利活用が可能と考えられるものである。「**第２章　ＡＩネットワーク化の進展に伴い形成されるエコシステムの展望**」についても、将来的に AI の利活用があらゆる産業活動・社会生活面において進展することを見据えて行っているものである。

このように AI はその利活用の範囲が広範に及びうるものであるが、他方で「**はじめに**」で述べたように、AI に対する倫理的な不安を有しているのは、AI 開発者・サービスプロバイダーのみならず、AI に関する専門的知識が比較的少ない最終利用者側にもあると考えられる。ICT の普及促進にあたり、インフラの整備のみならず、その利活用を進めることが課題とされてきたように、AI においてもその利活用の推進が重要となる。そこで、ビジネス利用者には様々な類型があるなかで、現時点において AI の利活用に意欲的に取り組んでいる例、他方 AI の利活用にあたり課題となった例についてヒアリングを行うことで、AI の利活用を進めるにあたり課題となっていることは何か、またその課題解決に必要な取組は何かについて検討することとした。

２． （株）三井住友フィナンシャルグループ（「SMBC グループにおけるデジタライゼーションの取組み」）

（１） ヒアリング概要 [84]

ア． AI 活用の目的、方向性

大きな AI 活用の目的は 3 点。①利便性。AI を活用してお客様に提供するサービス価値を最大化したい。②効率性。行内業務の効率化、自動化など。事務作業の自動化や不正取引・アンチマネーロンダリング等への適用で AI を活用していきたい。③収益性。一企業として収益性の観点からも AI を活用したい。

---

[84] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

イ． 金融業務・サービスへのAI活用事例

事例として、①AIを活用したコンタクトセンター支援、②チャットボット、③分析高度化、④企業の業況変化検知、⑤AI株式ポートフォリオ診断、及び⑥AIによるプログラム不具合の修正技術を紹介。その他、茲許、注視している技術として2点紹介。

(i) GAN（Generative Adversarial Networks）敵対的生成ネットワーク[85]

GANを使ってディープフェイクを実装している例。オバマ大統領の例のように、人間が見ても本物と見分けがつかないほど高精度な偽物を生成できるレベルに達しており、悪用への懸念が高まっていることから、更に高い倫理観が求められている。

(ii) 説明可能AI（ブラックボックス問題）

ブラックボックス化が取りざたされているが、解釈性と精度はトレードオフの関係にある。ブラックボックスのモデルの解釈性の向上と、逆に解釈性が高いモデルの精度の向上を図る動きが活発で、技術動向を注視する必要がある。

ブラックボックス化の本質を捉えた例として、米国陸軍で、敵と味方の戦車を識別するAIを作成した事例を紹介。テストでは高い精度を上げたが、本番では精度が低い結果であった。訓練データを調査したところ、味方の戦車の画像は晴れの日のものが多い一方、敵の戦車の画像は曇りの日に撮影されたものが多くばらつきがあったため、空の様子を見てAIが判断してしまっていたという事例。検証用データセットを用いた検証でも良い結果が得られたことから本番の新しい未知のデータが来た時にも相応の精度が出るだろうと考えてしまうが、その判断根拠がわからないと、学習者の意図とは全く異なる要素から学習がなされ期待した結果が得られない事態が発生しうる。解釈性に乏しいモデルを活用するリスクをよく表現した事例。

ウ． 課題、ガイドライン整備状況等

AI特有のリスクなどをまとめてAI導入ガイドラインを早期に策定。リスクは①回答精度、②ブラックボックス、③AIエンジン毎の特性、④データのバイアスの4つに収斂。

AIの導入フロー、①企画、②学習、③導入後の活用の各段階でこれらのリスクをしっかり認識したうえで対策を明示。

エ． AI導入ガイドラインの構成

前段でAIの定義やリスク、導入時の留意点について記載。後段でAI導入のフローに沿

[85] 生成器（識別器が誤認識するようなデータを生成）と識別器（学習用データと生成データを識別）の2つのAIモデルを競わせて訓練し、学習データと似たような性質をもつ出力を生成するフレームワーク。

って企画からリリースまで、運用の各段階に分け、ポイントを絞ってガイドライン化。さらに、倫理的に不適切な結果とならないよう留意が必要である点や、知的財産権との関連、AI開発時のベンダーとの契約上の留意点等を整理。なお、これらを纏めるにあたり本推進会議での「AI 利活用原則」の原案の段階から参照し、礎とした。

オ．　AI 活用の今後と課題

(i)あらゆる業務システムへの AI の浸透及び AI 間の連携

従来は用途別に独立して AI を導入してきたが、今後はあらゆる業務システムの中に AI が当たり前のように浸透してくる。例えば企業の業況変化を予測する AI、それをもとに与信を算定する AI、さらにそれをもとに適正な貸金額や提供すべき金融商品を提案する AI など、一連の業務の各所に AI が浸透することを想定。「AI 利活用原則」の中でも「連携の原則」として明示されている。

(ii)　体制面

企業として AI を積極的に活用していく上で導入体制は非常に重要なテーマ。技術面の情報収集や技術の検証を所管する担当部署と、業務を所管する部署を融合しながら推進しないと有効な AI 導入の議論にはならない。AI に限らず先進技術導入にあたっては、行内各部が協働で推進する強固な体制を構築することが肝要だということ。

カ．　今後の展望

(i)活用方針

早くから AI 活用に取り組み、ガイドラインも策定してきた。今後も AI に限らず、量子コンピュータ、AR、VR、音声認識など台頭する技術を早期に見極めて引き続き積極的に活用していく方針。

(ii)　推進体制（**オ**．(ii)のとおり）

(iii)　人材育成

AI や先進技術に明るい人材を更に強化・増員していかなければという課題意識。Web での学習プログラムの受講、日本ディープラーニング協会の認定資格の取得等を通じて、育成活動を実施。今後もグループ全体の AI の知識や企画力を高めるために継続して人材を育成。

（２）　議論

【AI 間の連携の課題と改善策】

Q. AI 同士の連携について、どのような事柄が今後の課題となってくるのか。

A. あるAIが出した答えを次のAIが入力として連携をしていくとなったときに、最終的なAIが出した答えの判断過程が広範にブラックボックス化されてしまい、単体でのAI導入よりもさらに解釈性が低下する可能性がある。その改善策としては、極力解釈性の高いAIを導入する、あるいは業務のフローとしてAIが出したものは一旦人が見るようにし、連携により誤った判断となるリスクを極小化するなど、対策方針を明確に決めていく必要があると考えている。

【解釈性と精度のトレードオフ、説明可能AIについて】

Q. 枯れた技術でもディープラーニングに近いものを出せるようになってきたということがよく報告されており、場合によってはディープラーニングとほぼ同程度の精度も軽いデータで出すことができ、かつブラックボックス化も若干緩和できるといったところも含め、トータルで見てどういう方向性なのか。

A. 解釈性のトレードオフについて、従来の解釈性が高いアルゴリズムは精度が高くない傾向があったためそれを引き上げていくという研究がなされている認識。なお、チャットボットの開発の時はディープラーニングを使っており、リカレントニュートラルネットワーク（RNN）というディープラーニングに分類されるアルゴリズムを使って作っている。これは行内利用目的に開発した経緯に加え、行員が確認した上で追加学習させるモデルとしており、解釈性よりも精度を追求する形とした。このようにケース・バイ・ケースでAIアルゴリズムを使い分けながら導入を推進している。

【テキスト情報についてのシステムの考え方】

Q. テキストで情報が入ってくるものがものすごく多く、テキスト書類の分析がどのくらい強力かというところが決定的にシステムの性能に影響を与えると思うが、テキスト書類の分析についてはどのようなことを考えて、どういう方向で考えているのか。

A. 既にテキスト分析基盤というものを導入し、例えば営業員とお客様との折衝記録（テキストデータ）の中から、お客様への有効な提案に資する情報を抽出するようなモデルを構築した。学習では、今までの膨大な折衝記録の中からいくつかの観点での正解データラベルを付けている。

【方言の音声認識の問題】

Q. コールセンターの件で、全国のお客様とのやり取りの中で、方言の音声認識の問題や、どういった苦労をされているのか。

A. コンタクトセンターの学習の点について、方言やお客様の電話の音声は、音声認識として精度が高くないのは当初課題としてあった。したがって、使い方の仕組みとして、お客様の声とオペレータが話している声のどちらもテキスト化しており、AIに入力するテキストはオペレータが話したクリアなものを選択できるようにした。

【外部データの活用について】

Q. 倫理規定に関連して、外部のデータや学習済みのモデルを使用するということについて、データの中身やその偏りが気になってくると思うが、こういった外部データを活用する際のデータの評価はしているのかどうか。しているのであれば、どういった仕組み、方針を取り揃えているのか。

A. クラウド上で多くの利用企業が画像をテキストに変換して修正した履歴から賢くなったAIモデルの代表例としてAI-OCRがあり、行内の少ないデータで学習させるよりも圧倒的に精度が高いことがわかっている。できればこのままそれを使いたいが、どういうデータを学習したのかが見えないので、それに対するリスクをどう捉えるかを検討している。但し、OCRの用途は、あくまでも画像をテキストに変換するのみで、必ず従業員が補正するフローになると考えており、そこまで大きな業務リスクはないとも考えられる。逆に、違う例として、与信判断に使うためのAIモデルを外から持ってくるとなると、リスクは異なる。用途に応じて外部のものを使うことが想定される。

【AIの出力とその透明性の課題】

C. これまではAIのエンジンの出力をスタンドアロンで使っていたものが、今後はエンベッデッドマシンラーニングという形で通常のシステムの中に組み込まれていくことも考えられる。通常のシステムの出力したデータが、どこの段階のAIによってつくられたデータを元にしているのかわからない世界が近いうちにやってくると思われる。それがまた、システム間で連携していくと、AIの出力がどの段階で出てきて、その出力が普通のシステムでどのように加工されていくのかといったプロセスが見えにくくなってくる。どのようにそのリスクをマネージしながら次のステップに行くべきかを重点的に考えている。

３． 東京都（「東京都ICT関連施策」）

（１） ヒアリング概要[86]

ア． AIチャットボット

- 福祉保健局＜受動喫煙防止対策問合せチャット＞

東京都受動喫煙防止条例の対象施設や条例内の用語の意味など、都の受動喫煙防止対策に関する質問回答を行うAIチャットボットを導入（2019年（平成31年）1月サービス開始）

- 水道局＜水滴くん相談室＞

[86] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

お客さまからの問合せに係るサービス向上を図るため、水道申込み手続や工事情報など、知りたい情報へのアクセスを支援する AI チャットボットを導入（2018 年（平成 30 年）7 月サービス開始）

イ． AI チャットボット総合窓口サービス（ワンストップ化）（2020 年度（令和２年度）展開予定）

すべてのチャットボットを統括する総合窓口となる AI チャットボット共通基盤を設け、利用者はその総合窓口に問合せをすることで最終的に目的の回答に辿り着ける機能を提供。

ウ． 総務事務改革

所管する事業に関わらず、すべての組織に共通して存在する内部管理的事務（給与・旅費，人事，共済，福利厚生，契約，会計，物品，文書）について、事務の集約化や ICT 技術の活用による効率化を推進。

- AI の活用

すべての組織に共通して存在する事務・作業において、AI の活用により、業務効率化・生産性向上を推進する。当面、問合せ対応、議事録の作成、文書・資料のチェック・校正の事務をターゲットとして取組を推進。

エ． RPA、AI-OCR、ETL[87]を用いたデータ化

デジタルデータの行政利用の促進を図るため、RPA 等の技術を活用し、都が既に所有する膨大な紙ベース等の帳票を利用可能なデータとし、これまでのデータ資産の有効活用を実施。

オ． AI 会議録作成支援

庁内での試行により、議事録の精度、使用頻度と導入コストを精査し、本格導入に向けた取組を実施。

カ． AI によるオペレータ業務支援

水道局お客さまセンターに AI を導入し、会話内容を認識して文書化・分析するとともに、回答候補などの情報を提供するなど、オペレータの支援を実施。（2020 年（令和 2 年）2 月から導入）加えて、AI により得られるビッグデータを業務改善などでの活用を検討。

（２） 議論

【AI の利用についての都民への周知について】

Q. 東京都は都民対応に対して AI を使っている、こういう性質を持つものであるとの説明

[87] Extract, Transform, Load の略。

をどのようにされているのか。

A. AI をどう使っているかという意味では、今都庁が使っている AI の状況は、利用されている方が AI だと分かる状態（チャットボットである、機械であると分かる状態）。ただ、AI の活用フェーズがより広範で汎用的な機能になるとしても、最後は人が関与することは重要なポイントで、自治体としてはフェイストゥフェイスが非常に重要と考えている。そういった意味で、今後、より広範で汎用的な AI の活用が広まるにつれ、そういった AI を使っていることを説明していかなければいけないという意味でいいご示唆をいただいた。

【AI の使い方について】

C. AI だから間違えることはなく 100%ということを期待されると、AI の利用が進まなくなるのではないか。

C. 100%ではなくてもやはり行政のシステムがそれでよくなるのだったらどんどんやってほしいと思う。平日の昼に役所に行くのはなかなか簡単ではない。これが AI だって分かっていて、何らかもしかしたらミスする可能性があるものだということがわかった人がこれを使えるっていう仕組みであれば、それだけでも多分凄く助かるという人はたくさんいると思うので、可能なところからスタートするのでもやってもらえるととても使用者としては嬉しい。

A. AI ですべてを代替していくと考えている訳ではなく、マルチチャネルの 1 つとして、色んなツール・メソッドとして活用していければよいと考えている。チャットボットを含めて AI ですべてを負担させないということが大事である今の段階では思っている。

４． ヤマハ（株）（「ヤマハの AI 歌声合成～美空ひばりをよみがえらせた取り組み～」）

（１） ヒアリング概要 [88]

ア． 経緯と前提

歌声合成について 20 年以上前から手掛けているが、昨年全く新しい AI の技術を取り入れ大きく進化した技術を開発。それを活用して美空ひばりさんの歌声を復活プロジェクトに協力。なお、同社としてボーカロイド、AI 技術に取り組み始めているが、まだ全社的に今後 AI をどういうふうにやり、AI 倫理をどういうふうに考えていこうかというところはまだ十分には議論ができてない状況。今回は歌声合成について、どのような思いで取り組んだかといったところを紹介するという前提。

---

[88] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

イ．　現行の VOCALOID と VOCALOID：AI との相違

現行の VOCALOID は収録済みの音を接続して合成するのに対し、VOCALOID : AI は予め大量の楽譜と歌声の対応関係をディープニューラルネットワークに学習させておくことで新曲の楽譜が入ってきた時にそれに対応する歌声をディープニューラルネットワークによって推定することができる仕組み。楽譜の文脈を見て歌ってくれるところ、人間が VOCALOID：AI に対して音楽的意図をリクエストできることが優れているところ。

ウ．　音楽的意図のリクエスト

秋元康氏から、30 年ぶりに復活する美空ひばりさんの空気感を演出してほしい、あるいは、おひとりおひとりに歌いかけるような歌を合成してほしい、とのリクエスト。こうしたリクエストを直接ボーカロイド AI に出したらやってくれるかというとそこまでではない。今回のシステムで人間の意図をリクエストする余地は以下 2 箇所。一つは合成の際に訓練に使った多くの美空ひばりさんの楽曲のうち、特にどの楽曲の時のスタイルで歌ってほしいかというリクエストを行う方法。もう一つはもう少し根本的なやり方で、そもそも学習に使うデータをある特徴を持ったものに制限して合成するという方法。

(i) 楽曲スタイルによるリクエスト

秋元氏のリクエストを踏まえ、結論として愛燦燦のような晩年の歌声、優しさとか豊かさっていうものが出ているような歌声になるべきなのではないかと解釈。今回の楽曲は、秋元氏及び我々の考え方など、ある種音楽的意図をもって合成。

(ii)　データ選択による雰囲気の再現

楽曲中の語りの部分を合成するにあたって、最終的にはひばりさんが息子さんのために取っておかれた貴重な音声をお借りし、このデータを元にして最終的に楽曲中の語りを合成。同じテキストでも言い方の雰囲気で、伝わるものの意味が変わる。何かデータを入力したらひばりさんの声が出せたというレベルに留まるのではなくて、ある程度秋元氏や我々の音楽的意図を持った選択を実施。

エ．　VOCALOID：AI について

楽譜の文脈を見て歌ってくれるが、裏を返すと楽譜の文脈しか見られない。例えば本物の歌手がやるように伴奏の雰囲気に合わせて歌い方を変えることは今のところできない。今回取り組んだような時代背景や企画の意図、30 年ぶりに本当に大事にしていらしたファンの皆さんの目の前に立って、ファンの皆さんたちにメッセージを届けるという企画の意図まではさすがに AI が自発的に理解することはできない。そこでその部分を間による音楽的意図のリクエスト、楽曲スタイルの選択、データの選択で秋元氏の意向・プロデュースの意向を反映しながら合成。

オ．　AI 歌声合成技術の伝え方と感じ方

NHK スペシャルでこの企画が放映された直後と、その後紅白歌合戦に出た直後で SNS における反応に違いがあった印象

NHK スペシャルの直後は、凄い・感動・面白い・素晴らしい・良いという非常にポジティブな感情が支配的だったが、紅白の直後は、怖い・気持ち悪い・殴る・気分が悪い、「冒涜」というようなかなりネガティブな反応が多かった。

カ．　何への「冒涜」を感じているのか

「冒涜」という言葉を使っている方は一体何への冒涜と感じているのかを分析。(i)生命の尊厳への冒涜、(ii)人格の尊厳への冒涜、(iii)芸術あるいは創造活動というものへの冒涜が考えられる。これに対して我々がどう考えているのか、あるいはなぜ NHK スペシャルではこういう反応があまり出なかったのか。

(i)　生命の尊厳

AI というものが凄く得体の知れないものが何か人間あるいは生命の真似をしているという印象から来てしまっているのかなという感覚。NHK スペシャルでは AI の仕組みをちゃんと伝えていたので、そこの印象は払拭できていたと思う。そもそも人工知能という言葉もそうだが、何だか生命の活動になぞらえた言葉を使ってしまいがちで、それで得体の知れない生命の模擬みたいなものの印象が出てしまっているのかなという点がもしかしたら気をつけないといけない。

(ii)　人格の尊厳

NHK スペシャルと紅白の一つの大きな違いは、聞く側・受け止める側に心の準備ができていたかどうかということ。心の準備というのは、今から AI 技術による再現をしますということ。ひばりさん本人ではないけれども、でもその気分になって楽しみましょうという心づもりができていたかどうかの違いではないか。

(iii)　芸術あるいは創造活動の尊厳

これは作り手の態度の問題、あるいは作り手と受け止め手が深い尊敬を共有しているかどうかという問題。今回のプロジェクトはそのつもりで実施。これがもし何かその AI というブラックボックスにひばりさんのデータを突っ込んだらひばりさんらしきものが出てきましたAIってすごいでしょ、なんか感動するでしょというやり方をしていたら、創造活動、人間の芸術活動を蔑ろにしているように見えたかもしれないし、自身もそういう違和感をもったかもしれない。今回の活動は、みんながひばりさんの音楽を少しでも理解したい、一歩でも彼女の音楽と同じものに近づいて、同じものを再現したいという思いでやってきた

ので少なくともその態度には冒涜と言われるようなものは何もなかったと我々は信じている。

キ.　我々が向かおうとする先

今回のプロジェクトは、人間が AI とともに音楽を作る、そういう未来に向かっていきたいという例のひとつ。

技術面及びその倫理的な線引きの面の両面で伝わり方が難しいところがあったが、今後こういう技術が本当にいい形で人に伝わっていく、使えるようになっていくということを目指してこれからの研究開発を続けていきたい。

AI というのが何か得体の知れないものとして認識されてしまっているという節があり、これが AI の良い側面の利用を妨げている節があると思っている。ネガティブな面、できること・できないことという面も含めて、先ずは正しい理解を皆でしていきたい。その上でいい利用を積極的にするための仕組み作りをやっていきたい。是非これを題材にして議論を進めていけだけたら嬉しい。

（２） 議論

【VOCALOID：AI の PR 方法】

Q. NHK スペシャルはドキュメンタリーで開発のプロセスを伝えられたので反感が少なかった、他方、紅白ではそれがなかったので反感が起きたのではないかという話があったが、今後 VOCALOID：AI を世に出していくにあたり、どういう PR を考えているのか。

A.ご存命の歌手による AI 合成であれば、ご本人が認めていれば、倫理的問題のクリアというのは比較的容易だと思う。そういった場合には楽譜の文脈に応じてその本人なりの歌い方もできるが、人間がそこに意図を入れ込むことによって作品として仕上げていくこともできる。それによって色々な人が、今自分では本物の歌手を雇えないような人が音楽を発信できるような、より良い音楽を発信できるようになったりする、もちろん全員の合意が取れた上でそういうところに向かっていけるといいと思っている。

【製造者・開発者の責任としてのサポートの在り方】

Q.（お客様が）ロボットや VOCALOID に対し、リアルの人やペット等に対してと同様の愛着を持ってきた場合に、事業者としては商品生産・サービス提供を終了できないと可能性もあると思う。今回特にこれについては商品開発まで進まないということだが、その可能性についてどう考えるか。

A. 我々がソフトウェアのサポートを終了して合成できなくなったら悲しむ方も確かにいらっしゃるかもしれない。開発者の責任というものが付いてくる可能性があるというのはご意見としてありがたい。

【人格の尊厳】

Q. 今回一番新しいのが、人格の尊厳に関わることも出てきたところではないかと思う。同社が目指しているように、感動を与えるということは、人の心を動かすということで、逆にいうと人の価値観を誘導することもできるかなと思う。特に故人の場合は、本人が出てきて違うと言えない。したがって、気持ち悪さの原点は、なんとなく社会的影響力が強すぎるものが勝手に誰かに作られて使われてしまうところにあり、そこが重要なポイントではないかと思う。それについての考えがあれば聞きたい。

A. おっしゃるとおりだと思う。オバマ大統領の声で不適当な事を喋らせるような例はたくさんあり、問題だとは思います。今はそれが開発者の倫理観にかかってしまっているようなところもある。しかし、技術的には止めようもないし、他方で、技術の開発はいい面もあるという意味で続けていきたい。できることできないことをどうやって使うべきかをちゃんとみんなで考えていくということをやるしかないのかなと思う。こういう新しい技術が出てきたからこそそれに基づく倫理観の形成っていうのはもしかしたら教育レベルでやるべきなのかもしれないのではないか。

【アカウンタビリティとの関係】

C. 「AI 開発ガイドライン」、「AI 利活用ガイドライン」との関係でいうと、アカウンタビリティ等については複数のステークホルダの皆さんに入っていただいて検討している。

【動画と音声について】

Q. AI 美空ひばりへの反応について、動画と音声で影響（反響）の違いがあったと考えるのか。

A. 音声だけでお客さんが感動してくれたのかと考えると多分そうではなく、NHK スペシャルのステージの上にひばりさんが本当に蘇ったかのようなことができたから、あれだけのエンターテインメント空間、つまり AI だと分かっている上でも感動ができたという面があるのではないか。動画があったおかげで音声もより引き立ったのではないか。一方で、AI が気持ち悪いとの感覚を持っている人は、映像を AI だと思っているのか否かも気になる。何が AI なのか区別がついている人に関しては、音声はこうだよね、動画はこうだったねっていう冷静な議論をしている。しかし、そうでない人は、もしかしたら何かよく分からない何かが浮かび上がってきたあの映像そのものによって AI という得体の知れないものが出てきたと感じている部分もあると思う。やはり映像の影響力というのは大きく、音声のみよりも映像の影響力が大きいというのは認めざるを得ないところはあって、だからより気を付けなきゃいけないところはあるかもしれないのではないか。

【音声の悪用防止について】

C. 例えば自分の音声について今後技術が発展し、歌として楽しく利用する場合はいいですが、例えば自分が意図しないように発声されてしまって、何か変なことに使われてしまう可能性も考えられるので、予防策、セキュリティ等も含めて技術開発をしていただけるとありがたい。

A. この音声はAIに使ってもらうつもりで残すぞと思っていなくても何らかの形で音声は使われてしまう可能性がある。動画共有サイトに上がっているちょっとした音声からその人の音声の特徴を抜き出すことも本当に可能になりつつあるので、音声を出す方に注意してほしいというのはおかしいけれども、意識として必要かもしれない。例えば、写真については意識している方も多く、ネットに自分の顔写真を出すとどう使われるか分からないから、意識をしましょうという意見もある。他方で、テクニカルな開発としては、ある音が合成音なのかそれとも本物の声なのかを見破るという研究はかなりやられている。そういう方向ももしかしたら、我々が同時に開発していくべきものなのかもしれないが、これは本当にいたちごっこであり、（GAN等を使って）見破るシステムを使って合成の方を訓練する方法も行われていたりして、いたちごっこで開発が進んでいる状況。

【リアルとフェイクの区別の問題】

C. 作られたものなのかリアルなものなのかの区別が分からないのが不安なので、AIの発展と共にこれはやはりそうやって作られたものだっていうことが分かることが一番ではないか、見破るなんてことはできないのではないか。

A. 基本的にはこれはAIで作ったものですよ、それを分かった上で楽しみましょうっていうことを明示するのが作り手側の責任と思っている。一方で、エンターテインメントとしてそれをあまりやりたくないという気持ちもある。例えば今回のライブを開催する前に、これはAIです、分かった上で聞いて下さいねというアナウンスを必ずやってからやりなさいみたいなレギュレーションができた時に、それを興ざめと感じる方もいるかもしれないと思う。テレビドラマや映画の場合、ひととおり終わってからこれはフィクションですと言っていますが、何かそのうまいところが狙えるといいと思う。

【声の権利について】

C. 歌声は、著作権のうち人格権なのか、隣接権の方になるのか。隣接権だとするとこれは相続の対象にもできるし、金銭的な価値を持つということにもなる。一方、これは美空ひばりそのものの人格（権）だとすると、それは美空ひばり以外にすることができない。そのあたりの切り分けを実はするということが第一の課題ではないか。また、AIとして残されたものが人格そのものとなって、どんどん学習して賢くなったりする。そう

いうような変化を許すか許さないか、これも一つ権利関係としては重要な課題ではないか。それをどこまで遺族等がコントロールできるのか。あるいはエンターテイメントとしてお金儲けに使って良いか等を整理するといいかなと思う。複雑な権利関係が絡み合いを起こしうるので、地図を作る非常に良いタイミングではないか。

５． 匿名

（１） ヒアリング概要

ア． セキュリティシステムの進化と AI

セキュリティシステムは、通信の大容量化やセンサ・カメラ等の技術進化を取り込み、ドローン活用や画像分析を扱う高度なレベルに進化。背景として、高精細画像（4K）のカメラが安価になってきていること、本当に使える AI が出てきたということ、及び 5G の開始。これを機会にセキュリティモデルを変えていく方向。

今までのビジネスモデルは、監視員がずっとモニターを見続けるという監視方法、又はハードディスクにためたカメラ画像を後から見る方法で、全部事後処理。それでは今までないようなソフトターゲットに代表される事件やトラックテロみたいな事件などでは手遅れとなるため、いかに予知予兆を事前に捉えるのか、というビジネスモデルが必要。さらに人手不足の問題もあることから AI を一生懸命取り組んでいる。また、全業種がお客様であり、契約先における様々なニーズを元に AI を一生懸命育てている。

イ． 事例①真報誤報判断の自動化

警報画像から AI が人物の有無を判定し、監視員の真報判断業務を支援。将来的に、真報判断の「支援」から「自動判断」へ発展を目指す。

業務の効率化であり、自主的に導入していこうということで推進。

ウ． 事例②X 線検査アシスト AI（X 線検査業務の AI による効率化・省人化）

経験を積んだ検査員が目視で行っている X 線検査について、AI によって業務を効率化することを目指した研究。

エ． 事例③不審行動（万引き）検知 AI（監視カメラによる不審者（万引き）検知システムの構築）

(i) モデル

万引き犯を見つけてほしいというお客様からのニーズで実施。モデルは、カメラを設置し、カメラから映った画像を AI で不審者として検知した場合に、そのスタッフとか店員の方または保安員の方のスマホに何番カメラに不審者がいるぞと通知し、それを受けてお声がけをすることで予防をするというモデル。

(ii) 社会実装に向けた課題

検知した不審者の情報を同業の複数店舗間で共有することは、顔認証することによって技術的にはできるしニーズもある。また、他業者でもそれを共有することによって、不審者のブラックリストを共有化したいというニーズはある。しかしながら、今のところ同一店舗の中でしかできないという問題があり、お客様から、同じ業者の中の複数店舗、または業種を跨っての共有ができればという要望が複数ある。

オ. 事例④新たなおもてなしサービス

(i) モデル

カメラ映像の AI 解析により、「困っている方」の動きを検知し、警備員に通知する「おもてなしサービス」の実証実験を新丸の内ビルディングで実施。

(ii) 社会実装に向けた課題

オープン空間であって、カメラ映像に映っている人たちをどうするのか、顔認証をやるのはどうするのか、という問題があることから、この実証実験を行った時は個人情報保護に基づいた手続をやっているが、これをもっとオープンにできるようにするにはどうしたらいいのかが課題。

カ. 事例⑤総務省 5G 総合実証試験 2017

(i) モデル

高所カメラによる広域監視、火災やそれに伴う交通渋滞の発生の検知と位置を検出することで、消防による最適な消火活動や自治体等による適切な避難誘導の通知等を可能にするもの。

(ii) 社会実装に向けた課題

オ. の取組は、ミクロの視点、個人情報の観点である一方、これはオープン空間のマクロの話、例えば映っている建物が本当にこれを映していいのかとか、走っている車種が写っているということで、技術要素的には可能だが、これをどう社会実装していくのかが課題。

（２） 議論

【課題に対する取組の進め方について】

C. 顔認証はなかなか難しい問題で、対応の方法には様々な考えがあるので、現場で課題に直面されている同社が問題提起をしたほうが、政府で議論する時もある種の事実になると思うのでよいのではないか、また、個人情報保護法改正のヒアリング等があるが、同社が話題を出していくのが一番届きやすいルートで、いろんな形で声を出していただ

くことが非常に重要ではないか。

【AI の利活用における人間の介在と業務の方向性について】

Q. AI だけで完結するということはなく、どこかに人間が介入する。最終的に AI のシステムを導入することで、同社全体で見たら省力化になるのか、それとも新たな対応の局面が増えるのか今後の見通しを教えてほしい。

A. 業務によって省力化と新たな局面の対応の両方がある。2400 か所にいる人間がそのまま順々に減っていくというよりも、その人間に社会インフラの監視とかいろんなマルチタスクができると考えており、そのため今のベースの作業をどんどん減らすことによってマルチタスクで様々拡大していくという考え方に近い。

６． 三部裕幸構成員（「AI 利活用の視点からみたリクナビ事例について」）

（１） ヒアリング概要 [89]

（前提）本推進会議事務局の要請に基づき、「AI 利活用の視点からみたリクナビ事例について」と題してお話しする。関係する各企業を批判するなどの意図は全くなく、本推進会議の活動にあたり視点となり得るポイントを共有したいだけであるので、ご了承頂きたい。

ア． AI ビジネスと従来ビジネスとの根本的な 4 つの違い

1 点目は、リスクはデータの収集の段階から始まるということ。

2 点目は、AI が帰納的で事前に結果を想定しづらいため、それによって生じる結果についても対応は必要だということ。

3 点目は、従来では困難なこと、人間の判断能力の代替までも AI が行うという点。

4 点目は、明治時代からの法律を含め現行法が想定できていないことは多く、AI がその埒外のことをすることがあり得る点。

これらによって、従来ビジネスの関係者は未体験な領域に入ることになるし、関係する法律やステークホルダを意識しづらくなる。また、法律や条例が未整備であるといったこともある。所轄官庁と調整しなければいけない問題点も多く生じるが、そこまで気が回りにくく対処できないという事態も生じる。

イ． AI の利活用視点からみたリクナビ事例のポイント

以上を踏まえつつ、リクナビ事例を見たときに、大事なのは次の 4 点だ。

1 点目は「法令・法的問題点、そして監督官庁について検討が必要」だということだ。

---

[89] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

同じ法令をベースにしていても、AIの場合には、従来とは違う問題点が起きてくるというのがここでのポイントだ。だからこそ、監督官庁との調整が必要となる。これは、「AIだからこそ問題になる」のでその点への意識が必要だということだ。

2点目は「倫理・レピュテーションリスクへの配慮が必要」だということだ。

3点目は、AIビジネス特有の「関連するステークホルダの広がりについて意識することが必要」だという点だ。

それらを踏まえたうえで最後の4点目として「社内の取組」が課題となるということだ。「ガバナンス」に関わる取組となるため、各社でどのような態勢を取るのがよいかを検討することがとても大切になる。

以下、この4つの視点で、リクナビ事例について検討する（下記**ウ**. ～**カ**.。4点目を3点目よりも先に述べる）。

ウ.　「法令・法的問題点、そして監督官庁について検討が必要」という点について

- リクナビ事例で個人情報保護委員会が最も問題視しているのは、就活生の同意を得ていなかったことではない。2019年（令和元年）8月の勧告の原因となった事実3つのうち2つは、安全管理措置が講じられていなかったことが理由になっており、その点が同意を得ていなかったことよりも先に論じられている。そのため、最も重要な問題とされたのは、組織体制や全社的な意識といったガバナンス体制だった。これがここでのポイントだ。
- また、東京労働局（厚生労働省）が、職業安定法に基づき指導をした。職業紹介事業者など一定の者が、その業務に関し、労働者の個人情報を収集・保管・使用するにあたっては、その業務の目的の達成に必要な範囲内で収集し、収集の目的の範囲内でこれを保管し、及び使用しなければならないと定めているのが職業安定法だ。そのため、個人情報保護法だけの問題ではない。これがここでのもう一つのポイントだ。
- そして、公正取引委員会が昨年末に新たな指針を公表した。個人情報の不当な取得・不当な利用があった場合には優越的地位の濫用として独占禁止法違反になる可能性がある。少なくとも今後はこのような点も考えなければならなくなる。
- 以上から、「リクナビ事例の問題点は個人情報保護法で求められる就活生の同意が取れていなかったことだ」と認識するのは間違いで、それだけの問題では済まない。個人情報保護法上もガバナンス体制が問題にされていること、そして、他の法も関わるということがポイントだ。
- 他の法ということでいうと、憲法も問題になり、また、理論的には民法の不法行為が成立する可能性もある。
- ガバナンスの本筋である会社法も問題になる。リクルートホールディングスという持株会社の子会社管理に問題があったのではないかといった問題提起をされるおそれは会社法上存在する。また、金融商品取引法も関係する。主要株主が同社の株式の売出し

をした。このようにガバナンス問題や市場に関わる法律も影響することを理解しておくことが必要だ。

- EU の GDPR（一般データ保護規則）も日本の議論にも影響する。2019 年（平成 31 年）1 月、いわゆる十分性認定を EU から日本が受けた。認定の維持のため、日本としては、個人情報保護に関して事業者に厳しい態度を示すであろうと予測される。
- AI ビジネスはほとんどすべての法律に関わる。リクナビ事例に即して考えても、個人情報保護法だけの問題ではなく、さらに、同じ法律が適用される場合にも、これまでのビジネスではありえない法的問題点が発生する、あるいは発生しやすいということがここでのポイントだ。

エ.　「倫理・レピュテーションリスクへの配慮が必要」という点について

- リクナビ問題は、個人の尊重という憲法の理念にも関わる問題だ。2019 年（令和元年）8 月、個人情報保護委員会は、「学生等の人生をも左右しうることから、その適正な取扱いについては重大な責務を負っていると認められる」と勧告している。個人情報保護法やそのガイドラインには、「人生をも左右しうる情報についてはこういう取扱いをしましょう」といったことが明示的に書かれているわけではないが、個人の尊重という理念にも関わるような大切な事項の取扱いが不適切な場合には、監督官庁からこのような追及を受けてしまうということだ。そのため、憲法上の理念を含め、AI 倫理という文脈で語られていることも、このように法律の解釈に取り込まれ、それに基づいてビジネスが適切か否か判断されることがあるので、留意することが必要だ。

オ.　「社内の取組」「ガバナンス」について

2019 年（令和元年）12 月の個人情報保護委員会の勧告によれば、リクルートキャリアらは、「組織体制を見直し、経営陣をはじめとして全社的に意識改革を行」うことが要求されている。顧客となった 35 社は、「組織的な法的検討を行い、必要な対応を行うこと」を含めた対応を要求されている。そして、「個人情報を取得する際は、商品等の内容をできる限り特定し、当該利用目的の通知又は公表を適切に行うこと」が求められている。単にプライバシーポリシーを公表していれば事足りるとか、「同意する」にクリックすれば事足りるということには必ずしもならない。

そのため、AI の製品・サービスの顧客となる企業も、AI を作る企業・使う企業の個人情報保護体制を検討する必要がある。「AI を作る企業・使う企業の個人情報保護体制を検討するプロセス」を、自社のビジネスプロセスの中に入れていくことが大切な時代になったということだ。

また、内部統制システムを会社法上整えなければならない大会社等でなくても、個人情報の取扱いを誤れば先程述べた独占禁止法など他の法律の違反となり得る時代に入った。AI を開発する企業や使う企業、そしてそのサービスを受ける企業も、データや AI

を取り扱う場合には、さまざまな法律との関係でガバナンス問題を考えなければいけない。これもここでのポイントだ。

カ.　「関連するステークホルダの広がりについて意識することが必要」という点について

リクナビ事例に即して検討すると、普通は就活生の情報を勝手に顧客に提供した点に注目しがちだが、それだけではない。大学、（上場会社であるリクルートホールディングスの）株主、監督官庁、リクルートグループの無関係の従業員や他の会社もステークホルダだ。また、顧客自体にもステークホルダが存在する。そしてマスコミも連日報道した。

このように見ると、ステークホルダの範囲はAIビジネスではとても広くなりがちだ。AIビジネスを始めるにあたっては、少なくともビジネスの初期的な段階で、各ステークホルダとの関係でどのような法的・倫理的・社会的問題があるかを検討するプロセスを踏むことが必須だ。

顧客も東京労働局や個人情報保護委員会から指導を受け、その事実が公表され世間から批判されている。そのため、リクルートキャリアにとって大切にしたかった顧客との関係にも悪影響が生じてしまう可能性が生じたことになる。

このように、AIビジネスならではのステークホルダの広がりもポイントになる。

キ.　まとめ

AIビジネスに関連する法律や倫理、ステークホルダには大きな広がりがある。従来ビジネスと同じようなつもりで進めると、リスクが実現してしまう可能性が高い。そのような事態に陥らないために、ステークホルダと法的・倫理的課題の検討を行うためのガバナンス体制を整えることが今後の課題として必要だ。企業ごとにビジネスやもともとの体制が異なることから、対処の仕方も大きく異なってくる。その点を踏まえつつ、今後の本推進会議での取組にあたって、実務上の工夫や課題をフィードバックすることで今後ますます寄与したい。

（2） 議論

【AIガイドラインの実装とガバナンスについてのコメント】

C. もともと「AIガイドライン」をどう実装していくかというところで会社のガバナンスにつながっていくと考えていたところであり、今回の説明に全く賛成。

【本件の情報の使い方自体についての考え方についての質疑応答】

Q.　本件の情報の使い方自体についてマイナスの評価が行われたかどうかということについてどうお考えか。就活生の立場としてはやはり内定辞退をするということを自由にし

たいという気持ちはあったと思う一方、企業はそれを予測したいというニーズもあったわけだが、そういう情報の使い方自体が問題だという認識がこのプロセスに出てきたか。

A. 最近では内定辞退セットというものが売り出されているが、リクナビ事例からも内定辞退セットからも、内定辞退に関する商品やサービスというのは相当大きな市場であると感じられる。ただ、明確なのは、内定は辞退してもかまわないのが大原則だという前提があるので市場になっているのだということだ。だからこそ、リクナビ DMP フォローを購入していた大企業にとっては、文字どおり切実な問題だったのだろう。

このような状況を踏まえたうえで、何を個人情報保護委員会が問題にしていたのかというと、就活生に対する説明が適切にされず、適切に同意を得ていなかったというプロセスの点、そして適切なプロセスを進めるうえでのガバナンスが効いていなかった点だ。言い換えると、辞退率データを第三者提供し利活用すること自体が絶対に許されないかどうかについては、個人情報保護委員会は明言していないと認識している。

そのうえで、一般論として AI を使う企業や開発する企業がリクナビ事例でどうすべきだったのかというのはなかなか難しいところがある。辞退率データの利活用が認められるかどうかについて私見を申し述べることは差し控えたい。

ただ、AI やデータの利活用ができる状況にあるが、社会的な合意が得られていない、というケースでは、どのようにアプローチしていけばよいかの検討が大切になる。

リクナビ事例に即して言えば、データを得る側と取られてしまう側、採用する側と就活する側とで利害やニーズが全く異なる問題といえ、社会的な合意は得られていない。このような問題については、対立するステークホルダの利益を考慮したり、あるいは法的・倫理的にどのように考えるのかという意識を社内であるいは社会的に醸成したりして、実現したいビジネスを社会が受け入れるかどうかの認識を深めておく必要があったのではないかと思う。それなしでビジネスを進めてしまうと、リクナビ事例のような問題は起こりやすくなると思う。

【個人情報についての理解についての質疑応答】

Q. 非常に危惧していることは、個人情報は個人に紐づいてしまう。仮にそれが推定情報であっても、嘘の情報であっても個人情報化してしまう。そういうような現実的な立て付けを皆さんが理解しているのかどうか非常に懸念していて、むしろそれがないために、個人情報ではないかのごとく扱ってしまったという現場の人の感覚があったのではないかというところに心配があって、これが他の会社でも重要なポイントになってくると思う。

A. おっしゃるとおりかと思う。というのも、個人情報保護委員会は、昨年（令和元年）12 月の勧告内容で、リクルートキャリアにあるデータだけでは個人情報にあたらないけれども、それを顧客に提供したら顧客にあるデータと合わせ技で特定の個人を識別でき

個人情報、個人データになるというものを、リクルートキャリアがわかっていながら就活生の同意なく顧客に提供したのが問題だと述べている。この点は、これまでも解釈上は論点になっていて、法律実務家はそのような論点があると把握していたポイントだ。個人情報保護委員会は、リクルートキャリアが就活生の同意なくデータを顧客に提供したことについて、「法の趣旨を潜脱」したと述べている。

また、個人の情報もハッシュ化すれば個人情報保護法上の個人情報にあたらないという判断をリクルートキャリアはしていたようだ。ハッシュ化は暗号化とさえ言えず特定の個人を識別できてしまう可能性があるので、同社の考え方は個人情報保護法に照らして通用しない見解であると考えられる。事実、個人情報保護委員会もリクルートキャリアの見解を誤った認識であると指摘している。

このような議論を受けて、例えば、相手方に出す情報が自社で個人データでなくても相手方で個人データになることが明らかな場合には、個人データの第三者提供の制限が適用されることが、昨年（令和元年）12 月に公表された「個人情報保護法　いわゆる 3 年ごと見直し　制度改正大綱」で明確化されている[90]。このように、AI だからこそ浮き彫りになる問題が今後現れると思う。AI やデータに関連する新たな問題点は今後ますます現れると思うので、問題意識を醸成していく努力が必要となる。

【教育・学習などのデータの利活用についての質疑応答】

Q.　大学が今後、学生のデータをきちんと分析して、ポートフォリオを作るといった取組を MIT や国立シンガポール大学でやっていて、その学生の市場価値を上げるために、大学がどのように貢献できるかをしている。例えばあなたの場合はポートフォリオだと、まだここら辺が履修できていないから、まずいのではないか、もう少し学習した方がレベルを上げられるといったことがある。企業から採用のときに履修証明を出すようにということがあると思うが、そういうことをどう考えているか、気づきの点があれば教えてほしい。

A.　日本に置き換えると多くの分析を要する点であり、直ちに回答をするのは差し控えるが、教育・学習などのデータの利活用が難しいという声をよく拝聴している。内閣府や文科省から学習支援技術や Society5.0 との関係での大学の取組についての委員会の委員

[90] 三部構成員のヒアリングの後に公表され、第 201 回国会において可決成立した個人情報の保護に関する法律等の一部を改正する法律（令和 2 年法律第 44 号）により個人情報の保護に関する法律が改正されることとなった。同改正の施行後は個人関連情報（生存する個人に関する情報であって、個人情報、仮名加工情報及び匿名加工情報のいずれにも該当しないもの）を第三者に提供しようとする場合であって、その第三者がこれを個人データとして取得することが想定されるときは、例外を除き、本人の同意（本人が識別される個人データとして取得することを認める旨）が得られていること（状況によってはさらに一定の事項）について、あらかじめ確認することをしないで、当該個人関連情報を当該第三者に提供してはならないこととなる。

を拝命しており、また大学で教鞭も取らせていただいた関係で、教育関係や大学の方々と接することが多いが、やはりデータの利活用には関心のある方が多いように思われる。自治体でも、いわゆる個人情報保護条例 2000 個問題をはじめ、データに関わる法制度の在り方を懸念する方もいる。どのように利活用を進めるかということに関して、有益な議論で公表できるものがあれば、本推進会議でも情報提供していきたい。

なお、そもそもデータ化されていないものをどうやってデータとして取り込んでいくのかについても検討が必要である。法律の問題ではないが、タブレットを児童・生徒・学生一人一人が 1 台ずつ使えるようにしていかないと、データの有益な利活用が難しいのではないかということは議論になっている。今後も関心を持ち続けてフィードバックしたい。

７． とりまとめ

（１） AI ビジネス利用者による AI 利活用原則等及びガイドラインの策定支援

AI 利活用原則等及びガイドラインを策定している AI ビジネス利用者は一部となっている [91]。企業等がその事業活動に AI を利活用する目的は様々であるが、AI 利活用原則等を策定することは、AI 開発者・サービス提供者との関係では、当該企業等の AI 利活用の基本的な考え方が明確となることで AI システム・サービス開発を協働して進めやすくなるという利点がある。また、組織内利用の場合であっても対従業員等との関係でも安心・安全に AI を活用していること、さらには最も重要なものとして AI を利活用して事業として顧客に製品・サービスを提供する場合には、対顧客との関係で、顧客が安心・安全に製品・サービスの提供を受けているという信頼を得るための重要なメッセージとなる。本推進会議では、今後 AI ビジネス利用者での AI 原則等の策定の意欲的な取組をフォローするとともに、そうした先進的な事例の紹介とともに、引き続き「AI 利活用ガイドライン」を周知すること等を通じて、AI ビジネス利用者による AI 利活用原則等及びガイドラインの策定を支援していくことが必要であると考えられる。

（２） 具体的事例を通じた AI 利活用の各原則についての考え方の整理・蓄積

本ヒアリングでは、AI 利活用の具体的取組を通じて AI 利活用原則の各原則についての考え方についても議論が行われた。本ヒアリングではヒアリング対象分野が限られていたこともあり網羅的な議論とはなっていないが、こうした具体的事例での原則についての考え方を整理し蓄積していくことは、今後企業において、AI 利活用の判断をする際の参考に

---

[91] 前掲、脚注 70 総務省情報通信政策研究所情報通信法学研究会 AI 分科会 2019 年度（令和元年度）第 1 回会合における新保史生構成員（慶應義塾大学総合政策学部教授）発表資料「AI 原則は機能するか？」<https://www.soumu.go.jp/main_content/000660996.pdf> 6 頁以下参照。

なるものと考えられる。こうした観点から引き続き、本ヒアリングでは対象とならなかった分野も含め、意欲的な取組についてヒアリングを行い、具体的な事例における各原則ついての具体的な考え方の整理・蓄積を行っていくことが必要であると考えられる。

（３） 「AI 利活用ベストプラクティス」の策定について（「第３章９．（４）と同旨」）

第３章９．（４）では、AI 開発者・サービス提供者の視点から記載しているが、AI ビジネス利用者による AI 利活用原則等及びガイドラインの策定は、そもそも企業における AI 利活用の取組そのものが進んでいくことと「鶏と卵」の関係にあると考えられる。こうした観点からも AI 利活用の取組の参考となる「AI 利活用ベストプラクティス」の策定に取り組んでいく必要があると考えられる。

（４） AI 利活用に必要な制度的課題のフォロー

AI を利活用して製品・サービスを提供していくにあたり、例えば声に関する著作権など、既存の制度との関係についてあらためて整理・検討する必要がでてくる場合が想定される。こうした制度的課題については、情報通信法学研究会 AI 分科会や他の関係機関等とも連携を図ることなどにより、本推進会議においても引き続きフォローしていく必要があると考えられる。

（５） AI 利活用ビジネスのガバナンスの重要性

AI ビジネスに関連する法律や倫理、ステークホルダには大きな広がりがあり、従来ビジネスと同じようなつもりで進めると、リスクが実現してしまう可能性が高い。そのような事態に陥らないために、ステークホルダと法的・倫理的課題の検討を行うためのガバナンス体制を整えることが今後の課題として必要となると考えられる。また、ガバナンス体制の検討にあたっては、企業ごとにビジネスやもともとの体制が異なることから、対処の仕方も大きく異なってくるものと考えられる[92]。その点も踏まえつつ、AI 利活用ビジネスのガバナンス体制について、本推進会議として引き続きフォローし、研究していくことが必要と考えられる。

（６） その他

本章においても上述のヒアリング概要、議論のとおり、**第３章**同様「安心・安全で信頼性のある AI の社会実装」の課題について多岐にわたるものが提起されている。こうして提起された課題については、他の関係団体等とも連携しつつ、必要に応じ引き続き検討を行っていくものと考えられる。

---

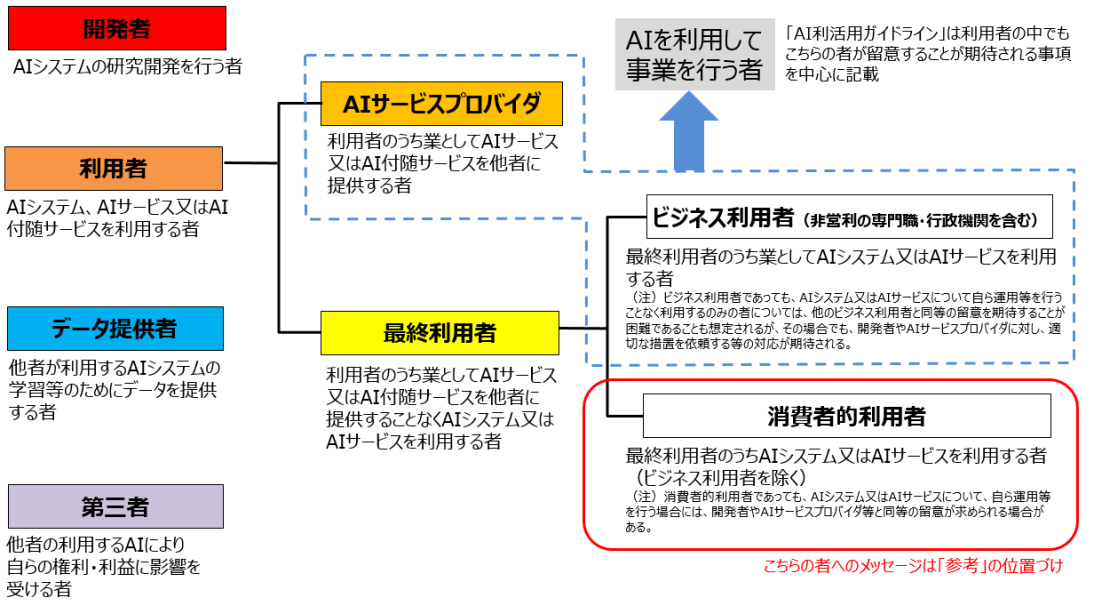
[92] 本章６．（１）**キ**．参照。

第５章 消費者的利用者に関する取組

１． 議論の出発点

AI 利活用ガイドラインでは、本ガイドラインの柱となる AI 利活用原則として整理した 10 原則について、AI サービスプロバイダー、ビジネス利用者及びデータ提供者が留意すべき事項について解説がされる一方、消費者的利用者については、留意することが望ましい事項について参考として併せて記載するものとなっている。

関係する主体の整理



（注）同一の個人・事業者が複数の主体に該当する場合がある。

今後、「安心・安全で信頼性のある AI の社会実装」を積極的に推進していくという観点からは、消費者的利用者が安心して AI を利活用し、その便益を享受できる取組が必要となってくる。この点については、『報告書 2019』においても、「**第３章　今後の課題**」において、「**１．ＡＩの開発及び利活用の促進並びにＡＩネットワーク化の健全な進展に関する事項**」の「**（１）　ＡＩ開発ガイドライン案及びＡＩ利活用ガイドラインの周知・展開**」において、「消費者的利用者向けにも分かりやすいメッセージを発信することが重要であり、AI 利活用ガイドラインにおける記載等をもとに、『ハンドブック』や『マニュアル』などリテラシー教材（利用者の手引き）を作成し、それらに基づいてワークショップを実施すること等についても検討することが望ましい。」との記載がある。

また、消費者的利用者のなかでも、高齢者・障害者が AI を利活用することにより、加齢あるいは障害を有することに伴う不便を解消することで、誰もが等しく自己実現を図れるようにすることは、人間中心の AI 社会を実現する上で最も重要な取組の一つであると考え

られる。こうした観点から、高齢者・障害者にとっての安心・安全で信頼性のある AI の社会実装の取組について関係者からヒアリングを行うこととした。

２． 消費者的利用者に関する取組

（１） 消費者の AI への関わりに関する政府及び消費者庁での議論

政府としては、「消費者基本計画」[93]の中で、「今期消費者基本計画における消費者政策の基本的方向」の 1 つとして、消費者が主役となる社会を実現するための AI をはじめとした革新的な技術との関わりについて以下を打ち出している：

「ICT の高度化によって、AI、IoT、ビッグデータ、ロボットの活用など、革新的な技術が発展しつつあり、我が国では官民挙げてこれら技術を最大限活用して Society 5.0 の実現を加速することを目指している。Society5.0 は、サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する、人間中心の社会であり、その中で消費生活も大きく変化すると考えられる。技術革新がもたらす変化のスピードは急激であることから、消費生活への影響等も大きいと考えられ、①経済のデジタル化に伴う取引や決済の分野の急速な変化や、②ビッグデータの利活用環境の急速な変化等への対応を重点的かつ迅速に進めていく必要がある。」

また、本記載とも軌を一にして、消費者庁において「消費者のデジタル化への対応に関する検討会」[94]が開催されている。その検討項目の一つとして「消費者を取り巻く AI 等の現状とそれへの向き合い方」が掲げられており、「AI ワーキンググループ」（座長：本推進会議の小塚構成員）[95]が立ち上げられ、検討が開始されている。

（２） 消費者庁の議論に対する本推進会議としての連携

以上の消費者庁の取組に対し、本推進会議としても連携をとりながら対応してきており、こうした連携を通じて消費者的利用者が安心して AI を利活用し、その便益を享受できる取組を推進していくこととしている。

具体的には、本推進会議であった議論等を踏まえ、「AI 利活用ガイドライン」において消

---

93 消費者基本計画（2020 年（令和 2 年）3 月 31 日閣議決定）
<https://www.caa.go.jp/policies/policy/consumer_policy/basic_plan/>

94 消費者のデジタル化への対応に関する検討会
<https://www.caa.go.jp/policies/policy/consumer_policy/meeting_materials/review_meeting_003/>

95 消費者のデジタル化への対応に関する検討会 AI ワーキンググループ
<https://www.caa.go.jp/policies/policy/consumer_policy/meeting_materials/review_meeting_004/> なお、本ワーキンググループには本推進会議の事務局である総務省情報通信政策研究所調査研究部から高木幸一主任研究官（現（株）KDDI 総合研究所 フューチャーデザイン 2 部門 共創戦略 2 グループ シニアアナリスト）が構成員として参加。

費者的利用者が留意することが望ましい事項について参考として記載した内容の例に加え、前述したとおり、その具体化・明確化及び消費者的利用者向けのハンドブック等の必要性について言及している。

原則ごとの論点と消費者的利用者に望まれること（例）

原則	原則に対する論点
①適正利用	ア 適正な範囲・方法での利用
	イ 人間の判断の介在
	ウ 関係者間の協力
②適正学習	ア AIの学習等に用いるデータの質への留意
	イ 不正確又は不適切なデータの学習等によるAIのセキュリティの留意
③連携	ア 相互接続性と相互運用性への留意
	イ データ形式やプロトコル等の標準化への対応
	ウ AIネットワーク化により惹起・増幅される課題への留意
④安全	ア 人の生命・身体・財産への配慮
⑤セキュリティ	ア セキュリティ対策の実施
	イ セキュリティ対策のためのサービス提供等
	ウ AIの学習モデルに対するセキュリティ脆弱性への留意
⑥プライバシー	ア 最終利用者及び第三者のプライバシーの尊重
	イ パーソナルデータの収集・前処理・提供等におけるプライバシーの尊重
	ウ 自己等のプライバシー侵害への留意及びパーソナルデータ流出の防止
⑦尊厳・自律	ア 他者の尊厳と自律の尊重
	イ AIによる意思決定・感情の操作等への留意
	ウ AIと人間の脳・身体を連携する際の生命倫理等の議論の参照
	エ AIを利用したプロファイリングを行う場合における不利益への配慮
⑧公平性	ア AIの学習等に用いられるデータの代表性への留意
	イ 学習アルゴリズムによるバイアスへの留意
	ウ 人間の判断の介在（公平性の確保）
⑨透明性	ア AIの入出力等のログの記録・保存
	イ 説明可能性の確保
	ウ 行政機関が利用する際の透明性の確保
⑩アカウンタビリティ	ア アカウンタビリティを果たす努力
	イ AIに関する利用方針の通知・公表

最終判断をすることが適当とされる場合には、適切に判断ができるよう必要な能力及び知識を習得しておくこと

AIが不正確または不適切なデータを学習することにより脆弱性が生じるリスクに留意すること

事業者等※の情報をもとに、必要に応じて点検・アップデート及びセキュリティ対策を行うこと

過度に感情移入すること等により、特に秘匿性の高い情報をむやみにAIに与えることのないように

自らの情報が正しく利用されているかを意識し、必要に応じ事業者等※に確認すること

AIの判断結果に疑義を生じた場合には、必要に応じ事業者等※に問い合わせること

※AIサービスプロバイダ、ビジネス利用者をまとめて「事業者等」と記載。

「消費者的利用者向けのハンドブックが欲しい」「具体化、明確化が必要」との声。

（３） 消費者庁「AI ワーキンググループ」における議論の概要 [96]

同ワーキンググループにおいて、議論を行っていくに当たり以下のミッションが共有された：

「AI を賢く適切に使うことによって消費生活に新たな便益をもたらし、消費生活を豊かにすることを最終目標とし、そのために事業者にどんな情報を開示することが望まれ、消費者にはどんな知識を習得することが期待されているか、さらにそうした情報のやりとりを促す観点から政府等の立場はどのようなものであるかを検討すること」

この内容を踏まえ、AI 及び AI と消費者の関係、AI の課題等について整理、認識を共有した上で、消費者が身につけることが期待されること（特に、AI の機能や AI の収集するデータについてが中心）やとりまとめの方向性及びその際の留意事項について議論を行っている。また、以上の議論を踏まえた消費者向けのハンドブックを作成することもあわせて計画しており、そのターゲットや構成、編集方針、内容についても併せて議論を行っている。

96 「消費者のデジタル化への対応に関する検討会 AI ワーキンググループ　会議資料」<https://www.caa.go.jp/policies/policy/consumer_policy/meeting_materials/review_meeting_004/>

３． 高齢者・障害者に関する取組

（１） 近藤則子構成員及び若宮正子様[97]からのヒアリング（「高齢者とAIスピーカの未来」

ア． 近藤則子構成員からのヒアリング概要（「スマートスピーカー調査中間報告について」）[98] [99]

介護を孤独に、介護を過酷なものにしているのは老人の重度の障害であり、老テク研究会の活動として 30 年前からパソコン、スマートフォン、そして AI スピーカを楽しく使えると、介護が改善されるのではないかという問題意識のもとでその普及活動を実施。

AI スピーカの国別普及率を見ると、中国が 22%、米国が 20%に対し、日本では 2018 年度（平成 30 年度）末で約３％。スマートスピーカーについて、欧米で報告されているようなネットワーク上のトラブルについて、トラブルの経験はなく、危惧している人は皆無。これまでシニア向けのスマホ講座を実施しているが、今後、スマートスピーカーの利用支援講座等において、先行する欧米での事例なども紹介していく予定。

イ． 若宮正子様からのヒアリング概要（「我が家における AI スピーカの利用状況と AI スピーカについてのアンケート結果に関するご報告（中間報告）」）[100]

我が家における AI スピーカの利用状況に関して、AI スピーカは口頭で操作できる利便性を有することから、接触によって操作する他の家電との間で状況に応じて選択的に利用できることがメリット。

AI スピーカについてモニターができたことで、聞きそびれたり、見そびれたりした場合にスピーカ又はモニターから補完する形で情報が入ってくることで、難視聴という高齢者の課題を克服。

AI スピーカについてのアンケート結果のなかで、情報漏洩等についての危惧に関する設問について、「あり」との回答は皆無。隠れた操作ミスがあとからでてこないかという人為的なミスの方がネット上のセキュリティよりも懸念。

高齢者というと介護・見守りの対象とされやすいが、高齢者は介護や見守りの対象とされることは嫌と感じており、自立支援と言われると前向きになることから自立支援の役に立つ AI 機器の開発を要望。

高齢者への AI スピーカの普及にあたってはネット環境が重要であり、サービス付き高齢

---

97 NPO 法人ブロードバンドスクール協会理事

98 ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

99 「スマートスピーカー調査」の報告は 2020 年（令和 2 年）3 月に開催された「スマートエイジングフォーラム・電脳ひな祭り 2020」にて実施。同フォーラムの模様は次に掲げる URL のウェブサイトに所掲。
< https://www.youtube.com/watch?v=Vk0czALp21U >

100 ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

者住宅の建設のときにインターネット環境を整える支援措置があることを期待。また、ネット環境の設定にあたりいわゆるお助けマンの支援員を地域包括センターに登録しておいて、設定する際の環境調査などをやってもらえることを期待。さらに、家電や住宅設備機器等の操作性向上を期待。

ウ. 議論

【AI スピーカからの情報漏洩について】

C. AI スピーカからの情報漏洩について危惧する回答が皆無であったことを受け、実際に情報漏洩の危険性がどれくらいあるのかわからず、また経済的損害があったということも多く聞かれないなかで、リスクアセスメントを考えていくことを通じてどこまで使っていいかということがわかるといい。

C. 情報漏洩が家族など身の回りで起きやすい現状を鑑みた場合に、今後 AI スピーカのインテリジェンスに頼ることとなった場合には、プライバシーのみならず金融資産等様々な情報が漏洩される問題がでてくる懸念がある。

【AI スピーカ依存への懸念】

C. 現在のスマホへの依存から、今後 AI スピーカに依存してしまう懸念が考えられるが。

A. AI スピーカの利用の現状からすると、まだお近づきにもなっていない状況であり、まずはお近づきになってもらうことが重要。

【AI スピーカが様々な用途で利用される際のリスク】

C. AI スピーカの利点としては、音声で簡単に操作できるというインタフェイスが非常にいいことがポイントになっており、その観点から将来的には手を使わないで様々な操作が可能になることが予想されるが、例えば自動運転での AI スピーカによる操作の場合には操作を間違えると非常に大きな被害ができることも想定されることから、今後 AI スピーカが様々な用途で利用される際にリスクというものを考えていく必要がある。

（２） 田丸健三郎構成員からのヒアリング [101]

ア. 高齢者に関する取組の概要

(i)背景

日本の成人 20 代、30 代、40 代の標準語の認識精度はかなり高くなっているが、高齢者や地方の方の音声認識性がまだまだ低い状況。特に高齢社会において、段階にもよるが認知症の方の介護での会話を観察すると、当然のことながら認知症でない方の会話とは全く異

[101] Microsoft における取組及び AI データ活用コンソーシアムにおける取組についてヒアリングを実施。ヒアリング資料（抜粋）は <https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

なる。

(ii) AI データ活用コンソーシアムにおける取組例（高齢者の音声データ活用による高齢者とロボットの AI による会話促進）

介護人材が不足している中で、認知症を患った高齢者の方向けに会話ロボット、アンドロイド等を作ってなるべく会話を促すことで少しでも認知症の進みを抑えるという研究開発を ATR、阪大の先生と推進。高齢者の音声データは少ないため音声認識の精度が低い傾向にあり、高齢者に特化して音声データを収集し、高齢者と会話するロボットなどに活用。

イ. 障害者に関する取組の概要

(i)背景（日米の比較）

米国の場合、聴覚障害をもっている学生がいる大学では、音声認識をして字幕で表示し、授業を進めるという手法があっという間に広がり一般的。なまっていてもかなりの精度で音声認識。専門用語も事前に使用する資料と power point を読み込み、同じ発音でも対象となる授業で使う言葉と認識したうえで字幕として表示するところまで現在では可能。

他方、日本語では方言も含めまだ困難。これは圧倒的に日本語のデータボリュームに差があることに起因。日本国内に目を向けたときいかに日本国内におけるサービス・品質を向上できるのかという取組を AI データ活用コンソーシアムが 1 つの取組として実施。

(ii) AI データ活用コンソーシアムにおける取組例（コンピュータ操作に困難のある方に向けたデータ活用プロジェクト）

肢体不自由の方が IT 機器を使ってコミュニケーションをする際に、視線をもとに文字入力してコミュニケーションができるというもの。これまではアルゴリズムベースの実装ではなかなか品質を上げられなかった（一つの操作に時間がかかることが課題）。深層学習・AI を活用することによって過去の入力をもとに学習して視線で予測し、入力していくところまで含めてこの実現を目指している。障害者向けの支援を行うメーカーが参加する日本支援技術協会と AI データ活用コンソーシアムで協働して取り組んでおり、日本マイクロソフトも協力。具体的にはデータ収集、分析、モデル化、システム化まで一緒に実施。

(iii) Microsoft における取組事例①（瞳孔変動を用いた重度障害者用意思伝達システム PuCom）

全く体が動かない方に対して瞳孔で画像分析、モデル化をして判断、意思表示をサポートするという取組。

(iv) Microsoft における取組事例②（AI と人とのハイブリッド情報保障システム AI ミ

ミ）

聴覚に障がいのある方が、講演会や講義に一緒に参加できるための情報保障システム。

前述のとおり海外では音声認識はかなり品質が上がっているが、国内でも日本語の音声データを集めて品質を上げることによって聴覚に障がいを持った学生などが学びの場へ積極的に参加できるようにしようという取組。

ウ．　課題

日本が高齢社会の中で、特に地方の方々に対して AI を用いて一体どのような社会的なサービスや支援を提供できるのかということも今後非常に重要になってくる。データで見てみると、自然言語の分野では明らかではあるが、他の分野においても高齢者や地方の方に AI を活用できれば良いと思われる領域に反比例するようにデータが存在しない。したがってこういったデータの流通、取得においても真剣に考えていくことが重要。

４．　とりまとめ

（１）　消費者的利用者に関する取組

ア．　消費者庁との連携

消費者庁において消費者向けのハンドブック作成の検討が進められているところ、その動きをフォローしつつ、PR 等連携して進めていく事が有効と考えられる。

（２）　高齢者・障害者に関する取組

ア．　AI スピーカの活用方法等の周知活動への協力

AI スピーカのように、その利便性から既に実利用が進められている製品・サービスがあるなかで、我が国における普及率は必ずしも高くない。アンケート結果によれば、その理由としてネットワーク上のトラブルを危惧してというものが皆無である一方で人為的なミスの可能性が挙げられている。これを踏まえると、AI スピーカ等の利用は個人の選択であるものの、その活用方法等の周知が必要と考えられる。そこで、こうした周知活動に取り組む団体等に対し、本推進会議としても協力していくことは有益と考えられる。

イ．　先進的な取組事例等の周知等の推進

本ヒアリングにおいて紹介された事例は AI の利活用による高齢者・障害者の生活支援に関する先進的な取組として有益であり、こうした先進的な取組事例について引き続きヒアリング等を通じて収集するとともに、その周知等の情報発信に取り組むことが必要と考えられる。

第６章 セキュリティに関する取組

１． 議論の出発点

安心・安全で信頼性のある AI のための環境整備の一環として、技術面での方策の検討が必要と考えられる。例えば、品質の確保 [102]、説明可能性の向上、認証、セキュリティの確保など様々な取組が考えられる。

そのうち、セキュリティの確保に着目すると、AI 開発・利活用ガイドラインにおいては明示的に「セキュリティの原則」を打ち出しているが、その趣旨は、AI の開発及び利活用の促進や AI ネットワーク化の健全な進展のために、特に AI を使う人を守るセキュリティが重要との認識を踏まえたものとなっている。

他方で、「安心・安全で信頼性のある AI のための環境整備」という観点で AI とセキュリティの関係（以下「AI×セキュリティ」と書く）をあらためて考えた場合、以下の 4 つの視点が存在する [103]。

(a) Attack using AI（AI を利用した攻撃）

(b) Attack by AI（AI 自身による攻撃）

(c) Attack to AI（AI への攻撃）

(d) Measure using AI（AI を利用したセキュリティ対策）

このうち、AI 開発・利活用ガイドラインにおける「セキュリティの原則」は前述のとおり視点(c)にフォーカスしているため、本推進会議としては視点(c)を個別に深掘りすることも重要と考えられるが、「それぞれ深めるのも大事であるが、組み合わせることで研究が深まる。」[104]との意見もあるため、他の観点も考慮することが重要と考えられる。

以上を踏まえ、上記の観点から AI×セキュリティについてその対策の検討及び啓発活動を行う事業者団体である JNSA（日本ネットワークセキュリティ協会）よりヒアリングを実施し、AI×セキュリティに関する論点整理をすることとした。

---

[102] 機械学習に代表される AI のプロダクトに対する品質を確保する観点から、そのための指針を検討する QA4AI コンソーシアムが存在する。同コンソーシアムの指針では、分野共通の品質保証の枠組に加え、スマートスピーカー、産業用プロセス、自動運転等、具体事例にも触れており、第 4 章（ビジネス利用者に関する取組）の検討の参考になるとの意見もある。
< http://www.qa4ai.jp/ >

[103] 佐々木良一（東京電機大学研究推進社会連携センター顧問　客員教授）「AI とセキュリティ」<https://digitalforensic.jp/2018/09/18/column531/>

[104] 総務省サイバーセキュリティタスクフォース（第 20 回）資料 20-3「今後重点的に取り組むべき研究開発について」<https://www.soumu.go.jp/main_content/000666230.pdf>

２．　特定非営利活動法人 日本ネットワークセキュリティ協会（JNSA）

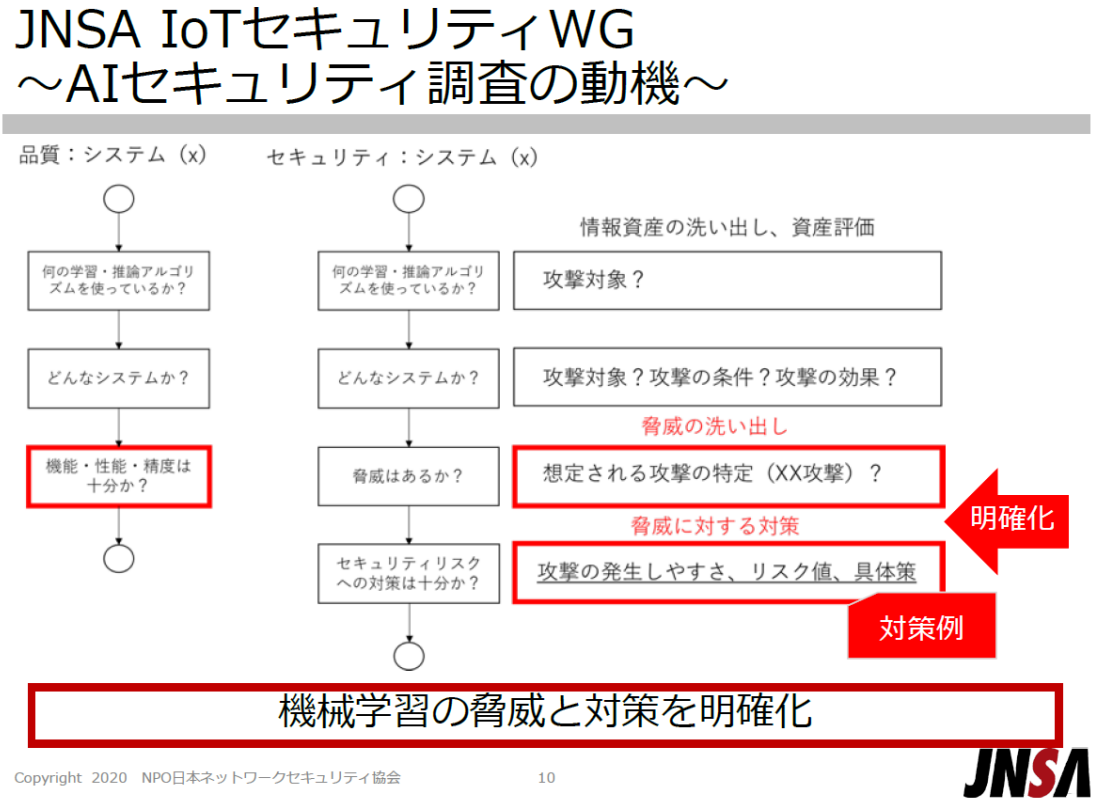
（１） ヒアリング概要 [105]

ア．　JNSA での取組

JNSA は 2000 年（平成 12 年）4 月にセキュリティ事業者の団体として発足し、現在は様々な企業のセキュリティ担当者も加わり、全 239 社が加盟する組織（2019 年（令和元年）10 月 25 日現在）になっている。安全なネットワーク社会を実現するため、ベンダー間の情報共有及び利用者向けの啓発・情報共有を、様々な部会等を通じて行っている。

イ．　「AI×セキュリティ」の全体構成と以下の議論

本章「１．**議論の出発点**」に記載のとおり、AI×セキュリティにはさまざまな視点があるが、 視点(b)を研究している人はあまり多くなく、それ以外の 3 点を整理することが多い。以下では、(c)AI への攻撃と(a)AI による攻撃について説明する。前者については、AI システムに対する特有の脅威・対策がないことから JNSA の IoT セキュリティ WG において、それらを明確にするために検討した内容について説明する（下図参照）。他方、後者については、海外のセキュリティカンファレンス等での発表を踏まえた研究の内容について説明する。



[105] ヒアリング資料（抜粋）は <https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

ウ.　(c) AI への攻撃について（分類と対策、方針）

AI、特に機械学習（モデル、システム等）への攻撃（脅威）については、大きく、以下の３つに分類される：

- 回避攻撃：人間には認識できないデータ入力により人間と機械学習の推論エンジンとで異なる認識を起こす攻撃
- 中毒攻撃：学習データへの不正データの入力により、学習モデルの境界を何らかの方法でシフトする攻撃、及び
- 移転攻撃：機械学習の推論エンジンへのデータ入出力または反応により元データなどの機密情報等を抽出する攻撃

AI特有の脅威

機械学習への攻撃

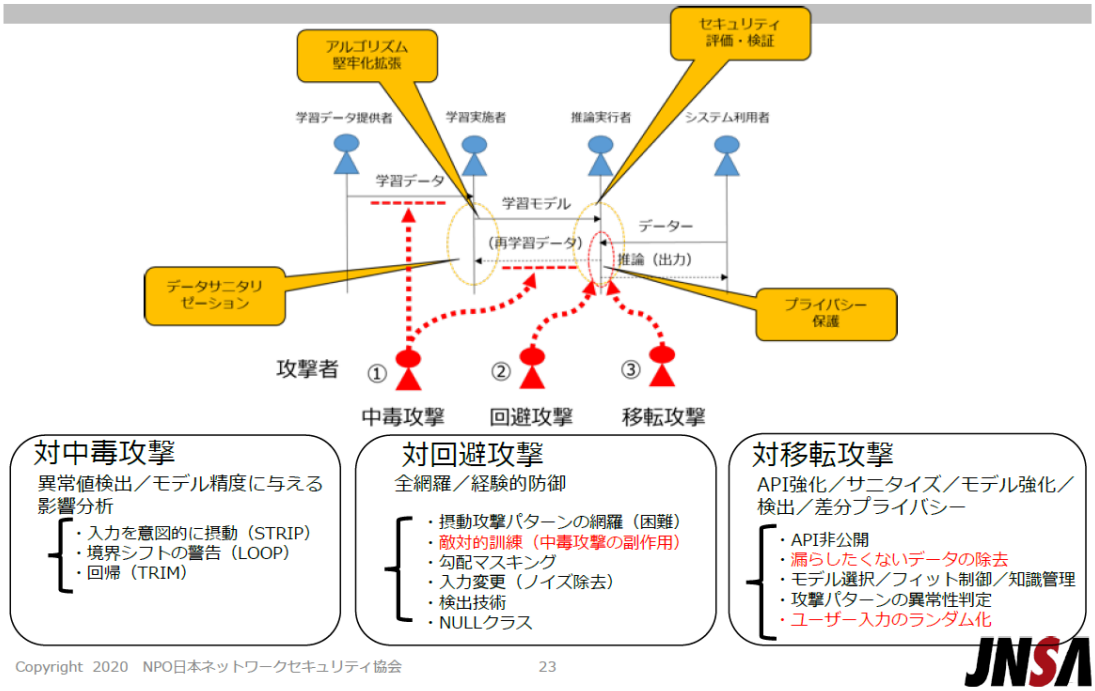
攻撃（脅威）	サブ分類	内容
回避攻撃 (Evasion Attacks)	・透明化攻撃 (Stealth) ・なりすまし攻撃 (Impersonate)	人間には認識できない「摂動を含んだデータ入力」により、人間と機械学習の推論エンジンとで異なる認識を起こす攻撃（画像、音声、文字等）
中毒攻撃 (Poisoning Attacks)	・可用性攻撃 (Availability) ・バックドア攻撃 (Backdoor)	学習データへの「不正データの入力（注入）」により、学習モデルの境界を何らかの方法によりシフトする攻撃 （機械学習のモデル境界を大量の不良データの注入により使用不能とする可用性攻撃と、少量の洗練されたデータ注入によりバックドアを生成するバックドア攻撃がある）
移転攻撃 (Inversion Attacks)	・プライバシー攻撃 (Privacy) ・メンバーシップ推論攻撃 （Membership）	機械学習の「推論エンジンへのデータ入出力または反応」によって、元データなどの機密情報（またはモデル自体）を抽出する攻撃 （メンバーシップ推論攻撃では、敵対者が手元データが相手のデータセットに含まれているかを探る攻撃）

※違う名称、分類もあるが、ここでは３つの攻撃まとめた

JNSA

これらの攻撃に対する対策について、論文等で紹介されている有効な手段を整理すると、中毒・回避攻撃に関しては比較的 AI に基づいた技術、つまり AI のアルゴリズム化を直していくような感じが望ましく、他方で移転攻撃のようにデータを盗むタイプはセキュリティ上の技術である匿名化技術を使うと良いことがわかってきている（詳細は下図参照）。

また、これに加えて、可能な限り状況に合わせて攻撃を制限したり、攻撃したのが誰かを見極めてそれにあった対策をしたりするということがポイントとなるとの説明があった。



エ.　(a)AI による攻撃について（種類、対策の取組）

AI（特に機械学習）を利用した攻撃について、いくつかの攻撃が検討されているが、ここでは、DeepLocker と DeepFake について、それぞれの内容と対策について説明。

DeepLocker は black hat[106] USA 2018 で紹介された深層学習モデルにマルウェアを埋め込んだ標的型攻撃の手法。暗号化マルウェアをアプリに内蔵させ、平時は良性アプリとして動作するが、その中で標的の情報を収集しつつ標的の有無を深層学習モデルで判定し、標的と認識した場合にマルウェア攻撃をするもの。これに対する具体的な対策はないとのこと。

他方で DeepFake はいわゆるフェイク動画を作るための技術。これについては様々な対策が行われており、例えば米国では大統領選挙を意識したフェイク動画を検知する取組が行われていたり、さらに同カリフォルニア州ではフェイク動画の作成者を提訴する権利を州民に認める州法[107]を定めていたりするとの説明があった。

106 <https://www.blackhat.com/>
107 Assembly Bill No.602:
<https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB602>

AIを利用した攻撃

攻撃名称	AIの用途	内容
DeepExploit	最適な攻撃手法の判断	深層強化学習を侵入テストに利用したPoC。 侵入対象のシステムから収集した情報（OS、製品名/バージョン等）を基に、システム侵入に成功する確率が最も高い攻撃手法を判断して侵入行為を実行。侵入に成功後、侵入したシステムを踏み台にし、内部のシステムに侵入を繰り返す。
DeepLocker	・標的の識別 ・マルウェアの秘匿	深層学習を標的型マルウェア攻撃に利用したPoC。 暗号化したマルウェアを内蔵し、平時は顔認証アプリやビデオアプリ等として振る舞いながら、Webカメラ/マイク経由で標的人物の情報を収集。標的人物を識別した場合、内蔵マルウェアを復号して攻撃を行う。アンチウィルスソフトに検知されずに標的のPC/スマートフォン等に入り込むことが可能。
tAIchi	マルウェアの自動生成	GANと強化学習をマルウェア生成に利用したPoC。 既知のマルウェアをGANと強化学習で変形させ、アンチウィルスソフトによる検知を回避するマルウェア（亜種）を自動生成。
Deepfake	・顔の入れ替え ・表情の再現	Autoencoder/decoderをフェイク動画の作成に利用した技術群。 オリジナル動画の顔部分を標的人物の顔に入れ替えることで、標的人物のフェイク動画を作成。

JNSA

（２） 議論

【視点(b)の扱いについて】

Q. (b) Attack by AI（AI 自身による攻撃）について研究をしている方が少ないとのことだが、実際にはどのような議論があるのか。

A. 特別の議論はないが、人間のリスク感度は高くなく想定が十分でないこともあるので、本視点に関する分類を行った佐々木教授が、ご自身の IT リスク学に関する専門的知見から、本件に関するリスクも想定すべきだろうということで分類の中に挙げられているのではないか。

【視点(c)：機械学習への攻撃と意図の有無について】

Q. 中毒・回避攻撃について、意図的な攻撃（悪用）と、意図的かわからない攻撃（誤用）についてセキュリティの面から考えられているか。（例えば、複数の人が偏った学習データを投入することで起きていることが、特定の人による攻撃なのか、結果として集団的な誤用を誘発するようなものだったのかなど）。

A. セキュリティの視点からは意図を含まないものは対象として扱うことはできず、品質の確保の問題になるという前提を示した上で、それらの区別は非常に難しい（例えば、上図「機械学習への脅威と対策」のプライバシー保護の箇所において、移転攻撃を狙ったものがある場合、何度も攻撃が続くようであれば通信トラブル（品質の問題）の可能

性もあるが、攻撃の回数やどういったサイトからの攻撃かを踏まえ、意図的だろうと判断できる場合にセキュリティ上の対策を講じるなど)。

Q. また、中毒攻撃において「不正データを入力する」とあるが、不正かどうかをどのように判断するのか。

A. 技術的には何らかの閾値を設けて異常値で検出する方法（それにより不正かどうかを判断する方法）が考えられるが、区別は非常に難しい。

【視点(a)：フェイク画像の作成とその善悪の定義について】

Q. Deepfake について、フェイクがそもそも悪いものなのか良いものなのかという定義が難しいと考えられるが、悪いフェイクとどうやって切り分けるのか。

A. 何かしらの被害者が出るフェイク動画は悪意のあるものだろうと考えるが、グレーゾーンのものについては線引きが難しいのではないか。

３． 補論

- 視点(d)については、技術的には国内外を問わず検討が進んでおり、視点(a)・(c)の技術同様に black hat 等のセキュリティのカンファレンスで常に情報がアップデートされている。我が国でも「AI 戦略 2019」[108]において「年々複雑化・巧妙化するサイバー攻撃に対し、『予防』『検知』『対処』の各フェーズにおいて、AI を活用した高効率かつ精緻な対策技術を確立」することが目標として掲げられており、情報通信研究機構等様々な研究機関で検討が進められている [109] [110]。
- 本分野の検討に対し総じて言えることとして、「セキュリティ技術者だけでなく、AI 関係者を交えて議論を進めていくこと」が重要との意見がある [111]。また、「今、人文社会科学的な研究がサイバーセキュリティの世界では比較的役に立つ要素になってきている。（中略）ソーシャルエンジニアリングや、e ディスカバリーのようないろいろな制度が、組織に絡んだり、心理学に絡んだりしている。」との意見もある [112]。

---

[108] <https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistratagy2019.pdf>

[109] 本推進会議 AI ガバナンス検討会（第３回）情報通信研究機構・高橋研究マネージャ講演資料

[110] 総務省「IoT・5G セキュリティ総合対策 プログレスレポート 2020」第Ⅲ章　総合対策の進捗状況と今後の取組（横断的施策）（１）研究開発の推進⑥AI を活用したサイバー攻撃検知・解析技術の研究開発　参照
< https://www.soumu.go.jp/main_content/000688845.pdf >

[111] 総務省サイバーセキュリティタスクフォース（第 20 回）資料 20-3「今後重点的に取り組むべき研究開発について」（再掲）

[112] 総務省サイバーセキュリティタスクフォース（第 20 回）議事要旨

４． とりまとめ

（１） AI への攻撃に対する対策の深化とそれ以外の論点への対応

- セキュリティの原則等を実用に資するものとするため、技術的には(c)AI に対する攻撃の分類等を踏まえ、可能な限り攻撃を制限すること、攻撃者が誰かを見極めて対策を強化することが重要と考えられる。
- また、AI に対する攻撃だけでなく、AI に関連するそれ以外の論点も考慮しながら議論を進めることが重要と考えられる。特に(b)AI による攻撃は詳細な検討が進んでいないが、AI 開発・利活用ガイドライン検討時に自律的に動作する AI や AGI についても考慮したのと同様に、リスクの１つとして捉えておく必要があると考えられる。

（２） 攻撃における意図の見極めの必要性

- (c)AI に対する攻撃、(a)AI を使った攻撃の双方について、何をもって不正・悪であるかを見極めること（攻撃していると判断すること）が困難であるため、その意図をどのように見抜いていくかが重要と考えられる。
- なお、上記参考として、総務省・プラットフォームサービス研究会では同最終報告書[113]の中で、偽情報（何らかの意図性を持った虚偽の情報）への対応の在り方として様々なレベル感があるとして、ファクトチェックの推進[114]や ICT リテラシー向上の推進など複数の論点について紹介している。また、AI により情報の流通のコントロールや削除等の対応を行う可能性があることから、そのための透明性・アカウンタビリティの確保については、本推進会議の「AI 開発ガイドライン」及び「AI 利活用ガイドライン」等を参照することが期待される旨が述べられている。

（３） マルチステークホルダによる学際的な議論の必要性

- 本分野については技術面だけでは解決できない問題が含まれるため、（セキュリティ）技術者だけの議論にとどめず、心理学・社会学等の知見も交えながら、学際的な議論を継続的に行っていくことが重要と考えられる。

---

113 報告書は次に掲げる URL のウェブサイトに所掲。
< https://www.soumu.go.jp/main_content/000668595.pdf >
また、同報告書を含むインターネット上のフェイクニュースや偽情報への対策についてまとめたものを以下に掲げる URL のウェブサイトに所掲。
< https://www.soumu.go.jp/main_sosiki/joho_tsusin/d_syohi/ihoyugai_05.html >

114 例えば Yahoo! Japan が個人に対する誹謗中傷等を内容とする投稿への対応と題し対策強化に加え、深層学習を用いた自然言語処理モデル（AI）の技術提供や検討会を設置した旨が報じられている。具体的に Yahoo! ニュースのコメントに対し、関係性の薄い情報として 1 日 2 万件の発信を削除したとの事例が挙げられている。
（Yahoo! Japan プレスリリースより：< https://about.yahoo.co.jp/pr/release/2020/06/01a/ >）

第７章 保険に関する取組

１． 議論の出発点

「安心・安全で信頼性のある AI の社会実装」を進めるためのセーフティネットとしての取組として、AI の特性を踏まえ、その開発・利活用に係る損害の補填等を目的とした保険の仕組みが重要になると考えられる。『報告書 2019』においても、「**第３章　今後の課題**」において、「**１．ＡＩの開発及び利活用の促進並びにＡＩネットワーク化の健全な進展に関する事項**」の「**（３）関係するステークホルダが取り組む環境整備に関する課題**」において、「AI の事故等に関する被害者の救済（保険等）及び被害発生の防止の在り方の検討」が、「**（６）利用者の利益の保護**」において「AI システムのリスクにより利用者等に被害が及ばないようにする方策の検討、リスクが顕在化した場合（事故等の発生時等）における責任の分配や利用者等を保護する仕組み（保険等）の在り方の検討」がそれぞれ掲げられている。こうした課題を踏まえ、AI に関する保険について先進的な取組をしている企業からヒアリングを行った。

２． 東京海上日動火災保険（株）（「AI の普及を支援する保険の機能」について）

（１） ヒアリング概要 [115]

ア． 分析手法

AI を取り巻くプレイヤーを「AI 事業者」、「AI 利用者」及び「AI 間接利用者」に分類した上で、プレイヤーごとの AI を取り巻く想定されるリスク例を整理。それを踏まえ AI 関連デバイスによる事故の責任主体に関する分析を実施。

イ． 既存の法規制のフレームワークの問題意識

AI 事業者と AI 利用者の責任について、AI 事業者においては、「AI が製造物に組み込まれる場合は、従来の製造物責任の枠組みで責任を整理することも可能だが、AI 固有の特性（学習能力とその行動の予見不可能性等）ゆえに、必ずしもすべての事故において十分に機能するとは限らないこと、ただし、AI 事業者としては、AI デバイスから生じる危険性を許容可能な程度に制御・軽減することが求められるため、全く責任を負わないとは限らない。他方、AI 利用者においては、「AI 利用者に過失が認められる場合には、不法行為責任が成立するが、自律的に行動する AI デバイスが引き起こす事態の予見可能性を一律に論じることは難しく、過失の有無を判断できない。」という実情から、既存の法規制のフレームワー

[115] ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

クだけでは責任の所在・分配を決めることは難しい、との問題意識を提示。

ウ.　事例①（自動運転の事故に対する保険商品）

こうした問題意識のもとで、具体的な取組事例として、自動運転の事故に対する保険商品として、「被害者救済費用等補償特約」を開発。運転者が操作する自動運転中の事故に関して、賠償責任の有無にかかわらず保険会社が保険金を支払うことができることとし、原因が不明確な自動車事故に対して保険金を支払い、迅速な被害者救済を図るもの。その上で、後々、賠償責任がメーカー側にあるということが決まれば、保険会社からメーカー側に求償する構成。

エ.　事例②（新築住宅の瑕疵に対する保険商品）

新築住宅の売買においては、住宅の品質確保の促進等に関する法律により、住宅の主要構造部分について 10 年間の瑕疵担保責任を負うこととされていることを踏まえ、住宅の売主が住宅瑕疵担保責任保険に加入する例について紹介。

オ.　AI を取り巻くリスクに対する保険のアプローチ案①被害者救済費用補償保険

上記ウ及びエの事例を踏まえ、AI を取り巻くリスクに対する保険のアプローチ案として 2 つのアプローチ案を提示。一つは、自動運転の例にみられるように、「被害者救済費用補償保険」として AI 利用者が提供した製品・サービスによって、AI 間接利用者が怪我などの被害を被った場合に、AI 利用者の責任の有無にかかわらず、AI 利用者が手配する責任保険の特約として AI 間接利用者の損害を補填するというもの。このアプローチは、ダイレクトに被害者を救済するという意味では、AI の社会的受容性を実現する一つの手段。ただし、これがあるから AI が普及するのかというと、そうではなく、AI の普及のためには、事故がおこらないようにするためにはどうするのかが必要。

カ.　AI を取り巻くリスクに対する保険のアプローチ案②AI 品質保証責任保険

住宅の品質の確保に関する事例を参考とし、AI 事業者が AI 利用者に対して、AI の品質を保証し、こうした品質を保証した責任に対して保険をあてがっていくという考え方。ただ、どういう品質を保証するのかが一番の問題であり、品質評価基準の策定とセットの仕組みになるもの。AI に関する品質評価基準や、品質を評価する監査機関の存在があって初めて適切な品質保証が行われ、その品質保証に対して保険が裏で支えるといったアプローチが有効な手段の一つ。

キ.　まとめ

「保険」は、AI の社会実装・普及に必要な「安全性」「社会的受容性」の実現に向けて、AI 事業者、利用者、間接利用者それぞれのリスクに対して、AI 固有の特性を反映させた補

償を提供。その機能は、AI を取り巻く責任を明確化する法的・技術的アプローチや、監査機関等によるガバナンス、品質評価のガイドライン等の整備と組み合わせることでより大きな効果を発揮する。

（２） 議論

【自動運転の事故に対する「被害者救済費用等補償特約」の枠組み】

Q. 被害者救済という観点から好ましい取組とのコメントがあるとともに、自動運転の車は非常に複雑な構造をしていて、事故原因が簡単にわからないことから、保険会社と自動車メーカーの間の情報の非対称性が大きく、保険会社にとって事後的な求償の観点から難しい問題があることから、その点についてどのような制度的担保で解決するかという質問があった。また、保険で事故の被害者である運転手の側で自動運転の車のソフトウェアのアップデートを忘れていた場合等、被害者側の過失の判断について難しい問題がでてくると思われることから、その点に関する将来的な方向性をどのように考えるのか。

A. 現時点で自動運転による事故の責任をクリアーカットにできる基準や方法はない一方、既存の自動車の通常の世界でも過失相殺については非常に難しい部分があるなかで、基本的に過去の判例に基づいて一定の基準があり、それに基づいて判断して保険金の支払いがされている。したがって、自動運転の分野においても事故の事例を積み重ねないことにはあらかじめ決めておくことは難しいこと、また、自動車メーカーとの情報の非対称性を踏まえると、保険会社としても完全に自動運転が実現する時までに、必要な専門的な知識を習得する必要がある。

【品質保証の意義】

Q. 品質保証をやることによって、AI 事業者側は後々保険料という形で返ってくるので、きちんとやらなければいけない。保険会社も、きちんとやったところに対しては保険金を払わなくていいし、きちんとやっていないところについては、被害者に保険金が出るが、当然翌年の保険料で AI 事業者に跳ね返ってくることになるので、では良いものを作りましょうというインセンティブが働くという提案であるという理解で良いのか。

A. そのとおりであり、AI の品質保証を対象とする保険の方は、まさに事業者側が品質を上げるというインセンティブを担保するためのもの。事業者側が品質を確保しなければ、ひいては AI そのものの普及にも影響してくるであろうというところまでイメージした上での提案である。

【品質保証責任保険のメリット】

Q. 品質保証責任保険でカバーされているということは、きちんと検査を受けて、品質も一定保証され、保険会社も保険でカバーしてくれる品質ですというマークが付くことに

よって、消費者もそういう商品をこぞって買うようになり、好循環になっていくというイメージでよいか。

A. そのとおりである。また、救済費用補償特約でも、例えば事故データがたまり、ある特定メーカーの自動運転車は事故率が高いことになると、そのメーカーの自動車を所有している方の自動車保険の保険料が上がるといったようなことも考えられる。現在もメーカーや車種によって自動車保険の保険料が異なるので、自動運転車においても、保険料に差をつけて事故防止のインセンティブを働かせるという考え方はある。ただ、もう少しダイレクトに品質向上のインセンティブを持たせることができるのが品質保証責任を保証する方。ダイレクトに保険料に反映させることよりも、お墨付きを付けるという観点では品質保証の方が直接的に品質を証明するような形になるが、品質評価をする機関等の仕組みがある前提なので、この点についてはまだ課題である。

【品質保証責任保険の制度設計について】

Q. 品質保証の水準はAIの提供事業者によって異なってくる可能性があるので、保険設計がこの場合は可能なのか。通常、できるだけ一律な条件でケースを集めて、事故率などを計算していくということがあると思うが、AIの提供は、製品とサービスにおいてSLA（Service Level Agreement）が違ってくると、保険が本当にできるのかという疑問が出てくるが。

A. 保険なので、同種のリスクを集めないと安定的な制度は作れないと考えており、そういった意味では、AIと一口で言ってもいろいろな用途があるので、用途や類型別に一定の母集団を作るのが望ましいと考えている。自由にSLAを設定されると保険会社としても困ってしまうので、品質評価基準やガイドラインのようなものがあり、そのガイドラインに沿った保証という形にするといったような工夫は必要と考えている。

【品質保証基準の担保】

Q. 住宅の場合だと、一定品質を保証しなさいという最低限のことが法律で決められていると思うが、AI品質保証責任保険は、AIに関しても何とか基準法のような、法的な最低限のものが必要かもしれないということなのか。

A. 必ずしも法律ではなくても良いと思っているが、業界基準のようなものは何かしら必要だと考えている。

【商品設計について】

Q. 商品設計について、救済費用補償保険は、確かに自動車を前提とすると、自動車の場合はもともとエンドユーザーに責任保険があることから、Third PartyとFirst Partyを被保険者にすることができるが、AI一般では、Third Partyの方は、むしろAI利用者が被保険者の保険で、費用補償の方はAI間接利用者が被保険者ということになって、

ちょっとそれをパッケージにするというのがイメージできないが、自動車以外でどういうことを考えているのか。また、品質保証責任保険については、これを実践していくと多分幾つかやり方があり、一つは船の世界でやっているように、格付けを付け、格付けで保険料を変えるというやり方。二つ目は、免責事由を入れて、品質評価基準に合致しないものは免責にして保険金は払いませんというやり方。この場合は、品質評価基準は相当低いというか、ミニマムレベルでやらないと、免責で保険金がゼロになってしまう。三つ目は、例えばAI開発のところで、ガバナンス、利用者側の原則を遵守しているかどうかということを保険金支払の条件にしていくというやり方があり、どう仕込んでいくかという問題がありそうだが、そこをどう考えているのか。

A.事業者が加入する生産物賠償責任保険のようなThird Party保険に費用特約を付けるというイメージであり、被保険者に利用者を入れるというよりも、あくまで企業側が責任を負えば、責任保険の構成で払うが、被害者を道義上救済する対応費用という形で、一旦費用を企業側に負担してもらうことを想定し、その費用を保険でカバーするというイメージ。

【因果関係の問題】

C. 責任の所在がはっきりしないと保険設計が難しいという説明があったが、AIの場合は、それ以上に因果関係の問題があるのではないか。特に民事責任の因果関係と、保険の因果関係、保険事故の因果関係が同じかどうかも含めて、実は日本の保険法は意外とはっきりしていないのではと思っているが、保険実務家の感覚として、そこはどう見ているのか。あわせて別に、因果関係と予見可能性なのか、相当因果関係なのかというところは相当難しいところだと思う。

A. 我々が取り扱っている賠償責任保険においても、民事上の法律で規定される賠償責任に沿った形で保険の適用を考えており、AIに関しても相当因果関係が立証されなければ、当然賠償責任は負わないという整理。予見可能性の問題なのか、相当因果関係の問題なのかというと、結局は両方だと思うが、AIに関して言えば、何かこういう形だったらメーカー側の責任だとか、企業側の責任だとかというような判断をする基準が定まらないので、いっそのこと契約上の責任にしてしまえばどうかというのがこの品質保証の話。一定の基準に満たなかったらこうしますという契約を結び、そうすることで相当因果関係などの論点を考慮せずに責任の所在を認めて、保険の対象にするというコンセプト。

３． 損害保険ジャパン（株）（「スマートファクトリーにおける保険活用について」）

（１） ヒアリング概要 [116]

ア． IoT/AI ソリューションと保険

IoT/AI ソリューションの浸透はまだまだ不十分で過渡期であるとの認識。その導入の阻害要因例、それについての対策（リスクの種類に応じた対策類型）の検討、対策の検討を踏まえた対策の実施としての①保険活用（リスクごとに対応する保険商品）、②IoT/AI 導入ノウハウの蓄積によるリスクの最小化を通じ、技術とファイナンスの組み合わせで、「IoT/AI ソリューション＋保険コスト＜既存コスト」を実現するとともに、データ、ノウハウの蓄積による新たなビジネスチャンスの創出を期待効果として取り組んでいくモデル。

イ． リスクごとに対応する保険商品

商品ごとにオーダーメイドでの設計になることを前提に、ビジネス連携起点（損害規模小発生頻度高）によるものとしての①性能リスクに対応した性能保証保険、②故障（予兆）リスクに対応した故障予兆費用保険、リスク規模起点（損害発生大発生頻度低）によるものとして③稼働停止リスクに対応した故障予兆利益保険、④サイバーリスクに対応したサイバー保険、⑤賠償（契約）リスクに対応した IoT/AI 賠償保険の 5 つ。

(i)性能保証保険

IoT/AI ベンダー企業がユーザー企業向けに、本番導入時に一定の性能（SLA：Service Level Agreement）を満たすことができない場合に IoT/AI ソリューション導入費用を保証する旨の約定を締結し、その約定を履行することによりベンダー企業に発生する損害を補償するもの。なお、SLA の内容や補償金額については、社会通念上妥当な範囲で設定する必要があるとしている。そのコンセプトは、通常 AI の性能不発揮、誤作動リスクはベンダー企業が抱えており、PoC（Proof of Concept）から本番環境への導入が進まない企業が多いことから、リスクシェア（保証）により導入を後押しする効果があるもの。

(ii) 故障予兆費用保険

故障予兆を検知したことによりユーザー企業が負担する追加費用（原因調査費用、駆けつけ費用、応急処置費用）を対象として、ユーザー企業が負担する費用を保険提供によりバックアップするもの。そのコンセプトは本番導入前に想定されるネガティブシナリオに対しての保険活用と位置づけられるもの。

(iii) 故障予兆利益保険

故障予兆を検知したことでユーザー企業が工場や設備ライン等を停止したことにより被

116 ヒアリング資料（抜粋）は
<https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000232.html>参照。

る利益損失（故障予兆費用保険では補償されないもの）を対象として、ユーザー企業が被る損害を保険提供によりバックアップするもの。そのコンセプトは、故障予兆により故障を未然に防ぎ、また保険により、稼働停止して復旧作業等を行う場合の利益損失を補填することで、導入前に比べユーザー企業の損失をトータルで軽減するもの。

(iv) サイバー保険

被保険者の抱える賠償リスク、事故等に急遽必要となる費用、被保険者の利益損害・営業継続費用を補償するとともに、セキュリティ管理会社等からの通知により、不正アクセス等のおそれが発見された時点で、不正アクセス等の有無を判断するために支出した外部機関調査委託費用やネットワークの遮断対応を外務に委託した場合に支出した費用を補償するもの。そのコンセプトは、ベンダー企業が被保険者となる場合には、Connected を前提とするクラウド提供やリモートメンテナンスを提供する場合は特にサイバーリスクが高まるため想定外の賠償リスクに対応するものであり、ユーザー企業が被保険者となる場合には、サイバー攻撃による利益損害、営業継続費用を補償するもの。

(v) IoT/AI 賠償保険

IoT/AI ベンダーが提供する IoT センサの不具合、または AI（例：故障予兆）の誤検知等に起因して発生したユーザー企業の設備、機械等の損害について、ベンダー企業が契約書上の賠償責任を負うことにより被る損害を補償するもの。そのコンセプトは、例えば故障予兆はその効果があるといっても 100%とは言えず、故障予兆の見逃しを原因としたユーザー企業の損害を、ベンダー企業側がリスクを保有することで、ベンダー企業にとってのサービスの拡充、またはユーザー企業にとっての予算の固定化につながるもの。

なお、紹介した 5 つの商品以外のリスクも考えられるものであり、この 5 つの保険に絞った形ではなく、新しいリスクとして保険化していきたいという趣旨のもとでの取組としての位置づけ。

ウ. 保険設計の留意点

保険を検討するうえで重要なポイントは 2 つ。1 つ目は IoT/AI ビジネスモデルの関係者の責任分担の整理であり、誰が誰に対してどんなリスクを負っているのか明確に整理をする必要があると考える。2 つ目は、そのリスクが保険転嫁すべきリスクか否かという点であり、各企業のリスクマネジメントに適した設計が必要であると考える。

（２） 議論

【リスクの発生確率の見積もり】

Q. リスクの発生確率はどのようにして見積もるのか。

A. 従来型のリスクではないと思っており、データというところが一つキーポイントとな

ってくるかと思っている。PoC によるデータを得られることを前提に、故障予兆の発生頻度を計算できると考えている。

【サイバー保険】

Q. サイバー攻撃の場合、層的攻撃などで、実際システムを使っている側が守りにくい状況で攻撃してくるケースもあり、それが事故に繋がってしまうというケースがあるが、その場合、どのぐらい攻撃されたら保険に入っている事業者が責任を負うことになるのかという辺りの判断基準はどのように考えているのか。

A.サイバー攻撃に限らず、賠償保険は、お客様の方で賠償責任を負うか負わないかという判断になるので、ベンダーが賠償責任を負うかどかは、ケース・バイ・ケースであり、場合によっては、裁判での判断となることもあり得る。

【保険の免責事由について】

Q. 保険の免責事由について、故意はあるかもしれないが、基本的には免責事由はサイバーリスクだけだという理解でいいのか。そして、その場合について、サイバー保険で手当てするという全体像になっているということなのか。

A. 一般的に故意などといったものはもちろん免責になっている。性能保証や故障予兆といったものについては、お客様から PoC でデータをもらい、それに基づいて設計を行うので、条件決めや、実際に違った環境で運営しているとなると、予め設定した閾値や条件が変わってしまうため、条件や免責の設定は個別に判断する。

【商品設計の体制について】

Q. 全般的な協業の流れで、ニーズの確認から入って打ち合わせをして、特に PoC が重要になるかと思うが、そうして保険契約に至るという一連の流れの中で、かなりいろいろな人が関与しないとできないと思っている。特に、性能保証保険のところは、契約書をどうしているのか。どういうチームで、どのような入り方でやって、全体図を作り出しているのか。

A. この商品を設計していくにあたって、ベンダー企業において実際に運用している方やユーザー企業の方に対し、例えば、実際に現場で困っている方の意見を聞き、どういったニーズがあり、それに適用可能な保険は何かというのは確認しながら進めていくという運用に対し、従来、保険会社としてはパッケージ型の商品を販売していたが、このようなニューリスクに対して専門家をふまえたチームで進めているのが現状。

４． とりまとめ

（１） AI に関する保険商品及びその導入事例の収集及び周知・共有

AI のリスク特性を踏まえ、損害の補填等を目的として、様々な保険商品が開発されることは、AI を利活用するにあたってのセーフティネットを整備する取組として重要と考えられる。そこで、本推進会議では今後とも AI のリスク特性に応じた保険商品が開発されることをフォローするとともに、保険活用の導入事例についても収集し周知・共有する取組が必要になると考えられる。

（２） AI の品質確保と保険

AI の品質を確保することにより AI の普及を図ることは重要であり、AI の品質確保の観点から、AI に関する品質評価基準や、品質を評価する監査機関等によるガバナンスと組み合わせた品質保証責任保険のアプローチは有効な手段と考えられる。その際、AI に関する品質評価は AI の用途や事業類型によって異なること、それに応じて品質評価の監査の手法・実現可能性も異なると考えられることから、AI に関する品質評価と保険に関する専門的見地からの検討をフォローしつつ、引き続き検討していくことが必要と考えられる。

（３） AI に関する保険の法的論点について

AI に関する保険を検討するにあたり、既存の法規制のフレームワークだけでは責任の所在・分配を決めることは難しいことも考えられる。こうした責任の所在・分配という民事法に係る論点、さらには新たな保険商品を開発するにあたって生じる法的な論点については、AI に関する保険の法的論点についての専門的見地からの検討をフォローしつつ、引き続き検討していくことが必要と考えられる。

結びに代えて

本報告書のとりまとめの考え方は「**はじめに**」で記したとおりであるが、あらためて、個別具体的かつ意欲的な取組等についてのヒアリング及び自由闊達な議論を踏まえてとりまとめられた点を強調しておきたい。これまで『報告書 2017』では AI 開発ガイドラインが、『報告書 2018』では AI 利活用原則案が、そして『報告書 2019』では AI 利活用原則案を踏まえた AI 利活用ガイドラインがそれぞれの報告書の中心としてとりまとめられている。本報告書は、これまでの報告書と比べると、その特徴を異にするものとなっているが、これは AI の開発・利活用に関する議論のフェーズが「原則の策定等」のみならず、「原則の策定等」を踏まえた「AI の社会実装」をいかに進めていくかに移ってきていることを背景とするものである。**第１章２**．において紹介した OECD のオブザーバトリ（AI に関する情報共有を進めるためのプラットフォーム）の運用開始等にみられるように、海外及び国際的な議論においても「AI の社会実装」を進めていくための動きが活発化している。本報告書の土台となった本推進会議議長によるヒアリングにおいて意欲的かつ貴重な取組等の報告、本推進会議の構成員との自由闊達な議論にご協力いただいた方々には、感謝してもしすぎることはない思いであり、あらためて御礼申し上げたい。

また、本報告書は、ガイドラインの策定等を目的としたこれまでの報告書と異なり、報告書の策定をもって一区切りとなるものではない。今後とも引き続き「安心・安全で信頼性のある AI の社会実装」に関する意欲的な取組等をヒアリングし、自由闊達に議論をし、「共有知」として整理していくことが、我が国における AI の社会実装を進めるために必要と考えられる。今回ヒアリングが叶わなかった方々には、今後のヒアリングを期待したい。また、今回のヒアリングでご協力いただいた方々におかれても今後の取組の進展等についてあらためて、ご報告・ご議論できる機会をいただけると幸いである。なお、**第１章１**．において、COVID-19 対策についての AI に関する取組の現時点における動向を調査、概観したものの、具体的な取組については今後ヒアリングを行い、議論していくべき課題と考えている[117]。

最後になるが、「安心・安全で信頼性のある AI の社会実装」を進めるにあたり、例えば、社会課題解決に向けた様々な領域の組織・団体との協働、AI 開発における開発者と利用者との協働、社内における AI 開発・利活用推進のための組織間の協働、ガバナンスに

---

117 社会的状況が刻々と変化していくなかで、重視すべき変化の一つは COVID-19 のように、それまで想定されていた範囲をはるかに超えた突発性でインパクトの大きいものであり、今回の報告書でも取り上げられているものである。今後もこうした突発性が高く、社会が急いで対応する変化があることも予想される。その場合は、たとえ、原則の一部を緩和してでも深刻な打撃を避けるための緊急対応をとることも求められることと考えられる。今後のことになるが、COVID-19 対応から得られたレッスンを一般化する形で、各原則について優先度や重要性について検討する必要があると考えられる、との意見があった。なお、関連する記述として『報告書 2019』45 ページを参照。

おける様々なステークホルダとの協働など、あらためて様々な「協働」が重要な視点になっていたと考えられる。本推進会議も様々な構成員の協働[118]により成り立っている会議体であり、本ヒアリングもご協力いただいた方々との協働により行われたものである。本推進会議として今後とも「協働」を重要な視点として、引き続き「安心・安全で信頼性のあるAIの社会実装」に向けた取組を進めていくこととしたい。

[118] 本推進会議において、構成員からは以下の点の検討の必要性について意見があった：
・AIが扱いうる常識の範囲が今後において拡大すれば、社会はそれ応じてAIによる自律的な判断に依存する範囲を拡大させるだろう。こうしてAIが人類の道具からパートナーへと変化してゆく可能性を踏まえ社会への影響を議論してはどうか。また、そのためにAI研究者にヒアリングしてはどうか。
・COVID-19対策を受け、短期的にはAIを含むICT技術の利用は推進される中で、プライバシーや意思決定の問題等が問い直されることになると考えられるため、本推進会議でもその先を見据えて論点を設けてはどうか。
・AI原則等の策定を受け、AIネットワークの適正かつ円滑な形成及び利用を目指し、その鍵を握るトラストを形成する方策を検討してはどうか。

（参考）報告書 2019 に掲げられている「今後の課題」

1. AI ネットワーク化の健全な進展に関する事項

(1) AI 開発ガイドライン及び AI 利活用ガイドラインの周知・展開：
AI 開発ガイドライン／AI 利活用ガイドラインの周知のためのシンポジウムの開催、国際的な枠組みにおける原則を実現するための詳説の周知等

(2) AI の開発及び利活用に関する原則・ガイドラインについての議論のフォローアップ
AI 開発／利活用原則・ガイドラインに関する国際的な議論のフォローアップと継続的な見直し

(3) 関係するステークホルダが取り組む環境整備に関する課題：
ステークホルダ間の協力・ベストプラクティスの共有、法制度等の在り方の検討等

(4) AI システム又は AI サービス相互間の円滑な連携の確保：
関係ステークホルダ間で共有することが期待される関連情報の範囲等の検討

(5) 競争的なエコシステムの確保：
関連する市場の動向の継続的注視

(6) 利用者の利益の保護：
利用者に対する開発者等からの自発的な情報提供の在り方の検討、利用者を保護する仕組み（保険等）の在り方の検討等

2. AI ネットワーク化が社会・経済にもたらす影響の評価に関する事項

(1) AI ネットワーク化が社会・経済にもたらす影響に関するシナリオ分析：
シナリオ分析の継続的な実施・国際的な共有等

(2) AI ネットワーク化の進展に伴う影響の評価指標及び豊かさや幸せに関する評価指標の設定：
指標の設定に向けた検討

(3) AI システムの利活用に関する社会的受容性の醸成：
社会における AI の利活用に関する受容度の継続的注視等

3. AI ネットワーク化が進展する社会における人間をめぐる課題に関する事項

(1) 人間と AI との関係の在り方に関する検討：
専門職（医師、弁護士、会計士等）と AI システムとの役割分担の在り方等の検討

(2) ステークホルダ間の関係の在り方に関する検討：
AI のリスクが顕在化した場合の責任の分配の在り等の検討

(3) セーフティネットの整備：
労働市場の動向の継続的注視、AI ネットワーク化の進展に伴う所得の再分配等格差防止の在り方の検討等