

TR-1091

AI活用における品質・信頼性・説明責任 に関する事例調査

Survey on quality, reliability and
accountability when using AI

第1.0版

2021年5月31日制定

一般社団法人

情報通信技術委員会

THE TELECOMMUNICATION TECHNOLOGY COMMITTEE

本書は、一般社団法人情報通信技術委員会が著作権を保有しています。内容の一部又は全部を一般社団法人情報通信技術委員会の許諾を得ることなく複製、転載、改変、転用及びネットワーク上での送信、配布を行うことを禁止します。

目次

<参考>	4
1 はじめに	5
1.1 背景	5
1.2 概括的状況	5
2 品質・信頼性・説明責任等活用原則の定義	6
3 AIの適用事例	9
3.1 通信設備における適用事例	9
3.2 電力設備における適用事例	10
3.3 モビリティサービスにおける適用事例	12
4 各所団体における検討状況	13
4.1 通信設備関連団体における検討状況	13
4.2 電力設備関連団体における検討状況	13
4.3 モビリティサービス関連団体における検討状況	14
5 GAP分析	16
5.1 通信設備関連団体の状況とのGAP分析	16
5.2 電力設備関連団体の状況とのGAP分析	17
5.3 モビリティサービス関連団体の状況とのGAP分析	17
6 まとめ	18
付録 <u>FG-AI4AD モリ－問題の調査の結果（2020年10月11-20日）</u>	19

<参考>

1. 国際勧告等との関連

関連なし。

2. 改版の履歴

版数	制定日	改版内容
第1.0版	2051年5月31日	制定

3. 技術レポート作成部門

第1.0版 : AI活用専門委員会

1 はじめに

1.1 背景

AIの活用により、社会課題の解決や企業業務等の改革に向けた検討や開発が各所で行われている。企業における働き方や業務等の改革にAIを適用する場合は、AIの適用範囲が企業の枠内に留まり因果関係の推定が可能なケースが多く、その結果、AIにより導出された結果の妥当性検証やその改善ループの再構成が容易であることが多い。

しかし、これを広く社会全般の課題解決等に適用する場合（例：医療診断等）は、AIが導出した結論に至る過程が十分に納得できる形で説明できないことや、品質、信頼性についても定量化が十分では無いとの懸念から、AI導入に否定的な意見も散見される。

上記のような状況を踏まえ、学会や内閣府、総務省等からAI活用の便益を高める一方でリスクを抑制するため、AI活用の社会原則やAI利活用原則に関する検討が行われ、AI活用において留意することをとりまとめた原則案が公開されている。

TTCのAI活用専門委員会は、①エッジ型アプリケーション高度化、②サービスデリバリー・運用自動化、③設備障害予測・保守効率化、④サイバーセキュリティ対策の通信に関連した領域におけるAI活用を目的に活動を行っている。

この活動の一環として、上記4領域においてAI活用の際の「品質、信頼性、説明責任」に関して、世の中でどのような検討や取組みが行われ、どのような事項が解決され、どのような事項が未解決な課題として残されているのかを調査し、TRとしてまとめた。

1.2 概括的状況

政府が提唱する、サイバー空間とフィジカル空間を高度に融合させたシステムを経済発展と社会的課題の解決に用いるSociety5.0において、膨大なビックデータを、AIを活用して解析・解決していくことが強く提案されている。そのような今後当たり前になるAI利活用のために、2019年3月に政府は政府統一のAI社会原則「人間中心のAI社会原則」（統合イノベーション戦略推進会議決定）¹を取りまとめている。この社会原則では、AIの発展に伴って我々がどのような社会の姿を目指すべきなのか、そしてその方向性（教育改革、社会実装、倫理）について規定している。

また、社会全体におけるAIネットワーク化の推進に向けた社会的・経済的・倫理的・法的課題を総合的に検討することを目的とした総務省「AI社会推進会議」²や、データの利用等に関する契約及びAI技術を利用するソフトウェアの開発・利用に関する契約の主な課題を整理した経産省「AI・データの利用に関する契約ガイドライン」³等のAIに関わる議論が進められている。

「AI社会推進会議」では、AIを利活用する人が不安や進展の阻害要因などの懸念を抑制するため、「安心・安全で信頼性のあるAIの社会実装」にフォーカスし、「国際的な議論のためAI開発ガイドライン案(報告書2017)」、「利活用に関する指針(AI利活用ガイドライン(報告書2019)」を策定している。また、様々な分野・領域で、開発者・利用者等の異なるステークホルダーへのヒアリングを行い、シナリオ分析結果を取りまとめている(報告書2020)。

AI利活用ガイドラインにおいては、適正利用の原則を始め、安全の原則、セキュリティ・プライバシーの原則、公平性・透明性・説明責任の原則などAI利活用の原則を10個に整理し、それぞれの原則について解説をとりまとめている。本ガイドラインは、開発者、ビジネス利用者や消費者などの参画を得てとりまとめら

¹ https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistratagy2019.pdf

² https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/index.html

³ <https://www.meti.go.jp/press/2019/12/20191209001/20191209001-1.pdf>

れており、様々な立場のステークホルダが議論をする際に共通に参照可能な文書となることを目的としている。

一方で、国民の生活を支えるインフラ（通信や電力など）に関しては、国民への安定的なサービス提供が必要不可欠である。通信においては、2020年より新しい移动通信システムである5G（第5世代移动通信システム）を使用したサービスが展開されており、サービスの高度化に伴い人々の生活や社会インフラとしての側面がますます強くなっている。この通信インフラにおいても、近年、メンテナンスや異常検知といった領域においてAIが活用され始めており、今後、ますます適用領域が拡大していくと想定される。本稿では、そのような重要なインフラ提供におけるAIの活用事例を紹介するとともに、今後ますます重要となってくる公平性、信頼性、説明責任の観点でどのような検討が行われているかについてとりまとめた。

AIに関する議論は、我が国だけでなく世界各国で始まっている。過去2年間に米国、カナダ、中国、デンマーク、EU委員会、フィンランド、フランス、インド、イタリア、メキシコ、北欧、シンガポール、韓国、スウェーデン、台湾、UAE等が、AIの使用と開発を促進するための戦略を発表している。これらは、人材育成、スキルと教育、公共および民間セクターの採用、倫理、基準と規制、データとデジタルインフラストラクチャーといったAI政策のさまざまな側面に焦点を合わせている⁴。

まず、米国では、米国科学技術委員会のネットワーク・情報技術研究開発（NITRD）にて、「米国人工知能研究開発戦略計画」（The national artificial intelligence research and development strategic plan）という標準とベンチマークを通しAI技術を測定評価するAIのトレーニングとテストの共用の公的データベース環境開発の戦略を公表している。また、2018年5月、米国国立科学財団（NSF）は、“米国産業のための人工知能宣言(Statement on Artificial Intelligence for American Industry)”を発表し、AI研究に毎年1億ドル以上の投資を行うことや労働者訓練の支援を行うことを表明している。一方、国防省研究機関DARPAの次世代AI開発プロジェクト”DARPA AI Next”が、今後のAI開発指針を示している。

欧州では、EUの研究開発プロジェクトHorizon2020(2014～2020年)において、7つの社会的課題を抽出（①健康 ②食料 ③エネルギー ④交通 ⑤気候 ⑥世界の変革 ⑦セキュリティ）し研究開発を実施しており、この中で様々なAI研究開発プロジェクトも推進している。AIの活用に関しては、欧州委員会は、産官学に加え消費者団体や労働組合、市民団体等、EU内外の複数のステークホルダを含めた、欧州のAIの将来を議論するアライアンス「European AI Alliance」を設立している。AIの高度専門家グループ（High-Level Expert Group on AI：AI HLEG）中心に「信頼できるAI」実現を目指すAI倫理ガイドライン「Ethics guidelines for trustworthy AI」を2019年4月に公表している、また、2020年2月「AI白書：欧州の優越性と信頼に向けたアプローチ」（WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust）を公表している。

2018年4月、欧州委員会（EC）は、Communication Artificial Intelligence for Europe を発表。ECはEUでAIの強化の必要性を認めた（再）立法とソフト・ローの取り組み、加盟国の調整、投資支援にコミットするもので、AIに15億ユーロを投資し、少なくとも5億ユーロの個人投資を誘致することを目指している⁵

2 品質・信頼性・説明責任等活用原則の定義

様々なインフラサービスにおいてAI利活用が始まっており、AIを活用した製品やサービスが、生活や社会へ大きな影響を及ぼすものと予想される。そのため、AIを活用したサービス等が増加した場合、AI技術に対する品質や信頼性がこれまで以上に求められてくる。現在、AI活用における品質や信頼性に関する定義は、複数の標準化や企業等で議論が開始され始めているものの、確定した概念・定義は存在しない。そこで、本章では議論中のAIに関わる品質の定義についてまとめる。

⁴ <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>

⁵ <https://medium.com/swlh/a-timeline-for-europes-ai-strategy-d2fc9f7bbcf1>

AIネットワーク社会推進会議（総務省）が公開する、AIの利用者が留意すべき10原則のうち、品質・信頼性に関する重要な定義としては、「適正学習」「公平性」、「透明性（説明可能性）」、「アカウントビリティ（説明責任）」の4つが挙げられており、定義としては以下の通りである。

1. 適正学習の原則

利用者及びデータ提供者は、AIシステムの学習等に用いるデータの質に留意する。

2. 公平性の原則

AIサービスプロバイダ、ビジネス利用者及びデータ提供者は、AIシステム又はAIサービスの判断によって個人が不当に差別されないよう配慮する。

3. 透明性の原則

AIサービスプロバイダ及びビジネス利用者は、AIシステム又はAIサービスの入出力の検証可能性及び判断結果の説明可能性に留意する。

4. アカウントビリティの原則

AIサービスプロバイダ及びビジネス利用者は、消費者的利用者及び間接利用者を含むステークホルダに対しアカウントビリティを果たすよう努める。

「適正学習」に関しては、利用者及びデータ提供者が、AIの特性を踏まえ、AIの学習等に用いるデータの質（正確性や完全性など）に留意することが必要となる。AIの判断は、事後的に精度が損われ、低下することが想定されるため、想定される侵害の生じる頻度、技術水準、精度を維持するためのコスト等を踏まえ、あらかじめ精度に関する基準を定めておくことが必要となる。

一方、「公平性」に関しては、システムの出力や振る舞いが、人種や民族、性別などの特性による差別、偏見、偏愛に相当するような不公平な偏りを示すことがないことが前提となっている。そのため、AIを活用したシステムにおいては、AIの開発と利用において常に公正さを持ち、AIが人間社会において不公平や格差をもたらす可能性があることを認識し、開発にあたって差別を行わないよう留意する必要がある。また、データ自体が満たすべき不変条件や整合性条件、学習対象となる判断の公平性、個人情報の有無など、データに対する制約を満たしているかなど、モデルを学習する訓練データにおいても不平等がないことが公平性には求められる。このため、特に以下の点について留意することが必要である。

- AIの学習等に用いられるデータへの留意

AIの学習に用いられているデータやデータに内在する社会的なバイアス等に留意が必要。

- アルゴリズムによるバイアスへの留意

アルゴリズムによりAI判断にバイアスが生じる可能性に留意が必要

- 判断への人間の介在

その判断を用いるか否かに関しては、人間の判断を介在させることが期待される

「透明性」に関しては、AIに利用されるデータの取得方法や使用方法、AIの動作結果の適切性を担保する仕組みなど、用途や状況に応じた適切な説明が得られなければならない。また、AIの提案を理解して判断するために、AIの利用・採用・運用について、関連するステークホルダに対して、必要に応じて開かれた対話の場を適切に持つ必要が生じる可能性もある。

- AIの入出力等のログの記録・保存

AIの入出力等の検証可能性を確保するため、入出力等のログを記録・保存とともに、保存目的やログの取得の頻度等にも考慮が必要

- 説明可能性の確保

AIが活用される社会的背景を踏まえて、どのような説明が必要かを分析し、それに必要な対応を講じることで、判断結果の説明可能性を担保する必要がある

「アカウントビリティ」に関しては、利用者に対して利活用に資する情報を提供するためにAIの技術的な特性を説明することが求められる。そのため、利用者等がAIの利活用について適切に認識できるように、AIに利用されるデータの取得方法や使用方法、AIの動作結果の適切性を担保する仕組みなど、状況に応じた適切な説明が得られなければならない。また、利用者の利益に重大な影響を及ぼす可能性がある場合には積極的に通知することが求められる。

一方で、AIを活用したサービスを想定すると、AIを研究開発する開発者、それを利用する利用者に分別される。また、利用者に関しては、開発されたAIをビジネス提供するために利用する「ビジネス提供者」とビジネス提供者が提供するサービスを利用する「消費者的利用者」に大別される。つまり、AIを活用するサービスでは、B2B2Cという関係になることが想定される(図1)。そのような関係性の中で、AIに関わる品質・信頼性に関する利害関係はサービス形態に左右され、どこが品質を担保すべきかという観点では、今後継続した議論が必要となるだろう。

このようなサービスプロバイダと消費者間におけるAIの品質・信頼性・説明責任に関するステークホルダ間の関係については、「報告書2019」[1]の中で定義されている。本稿では、3章に記載するユースケースにおいて、①(AIサービスプロバイダ-サービスプロバイダ間)と②(サービスプロバイダ-消費者的利用者間)間で取り扱われる品質・信頼性・説明責任の提供について考察していく。なお、AIの品質を考える上で、データ品質の考慮も必要不可欠である。これについては、サービスプロバイダが自ら必要なデータを収集し、利用しているケースが多いのが現状であるが、今後、サービスプロバイダがデータ提供者のデータを活用してサービスを提供するなどサービスの提供形態によってはこれも重要な検討事項になるだろう。

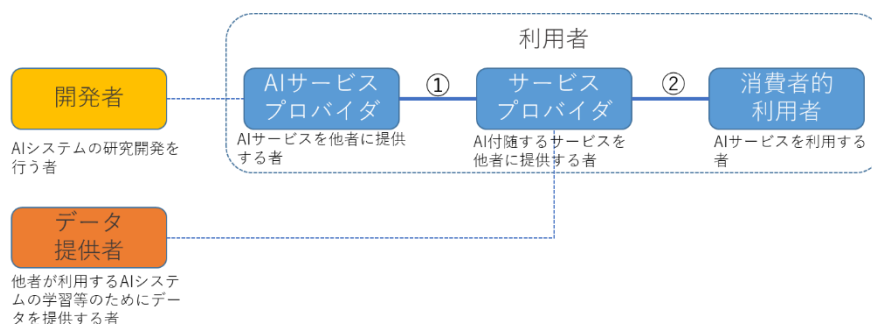


図1 AI提供の関係図

各企業や団体においても、AIの品質課題を解決するための活動が開始されている。

NECでは従来型のソフトウェア品質では対応できないAIシステムの品質を担保することを目的とした「NEC AI品質ガイドライン」⁶を策定している。同ガイドラインの特徴は、AIシステムのフェーズごとにチェック項目を設定するとともに、同社のAI開発の経験に基づきAIモデルに関する定量値を含むチェック項目設定などをガイドラインに盛り込んだことである。

NTTデータでは、これまでのユースケースやノウハウをまとめたAI開発のための品質アセスメントツールを開発し、2020年10月よりトライアルでサービス提供⁷を開始している。同品質アセスメントツールの特徴

⁶ https://jpn.nec.com/press/201912/20191210_02.html

⁷ <https://www.nttdata.com/jp/ja/news/release/2019/052100/>

は、プロジェクトの特性に応じたAIの精度だけでなく、活用によって発生しうる人的リスク、経済リスク、公平性、解釈性等について必要なリスクレベルを設定し、網羅的な品質検討を支援できることにある。

また、最近では富士通研究所と北海道大学から、「説明可能なAI」⁸の研究開発成果も報告されている。同技術を用いることにより、AIシステムの透明性と信頼性の担保の改善が期待されている。

3 AIの適用事例

3.1 通信設備における適用事例

通信設備でのAI活用に向けた検討事例として、以下があげられる。

1. 故障対応の自動化
2. 通信リソースの制御・オーケストレーションの自動化

1) 故障対応の自動化

通信設備の故障対応業務は、通常、故障の監視（異常の検知）、故障要因の特定、及び、回復措置実施の工程で行われるが、AIを活用し、各工程の自動化や効率化を目的とした検討が進んでいる。

通常、故障の監視では、CPU使用率や帯域利用率等、装置から得られる数値情報が事前に定めた閾値を超えた場合、あるいは、事前に定めたルールに該当するログが出現した場合等、異常を事前に定義しておくことで、検知する。しかし、検知した時点では、既に、サービスに何らかの影響を及ぼしている可能性があり、異常発生の手後での対応となっている。そのため、装置から得られる平常時のデータをAIに学習させ、平常時と違う兆候を早期に検知することで、サービス影響が発生する前の予防対処、もしくはサービス影響範囲ができる限り小さい段階での対処を目的にAIが導入されている⁹。

装置から得られる情報は膨大であり、個々の情報に対し、閾値やルールを定めるためには手間がかかる。事前に定めた閾値やルールでは検知できない想定外の異常（サイレント故障）が発生することもある。そこで、人間では判別が難しい正常時と異常時のデータ特性の違いをAIで識別することで、サイレント故障検知に活用した事例¹⁰や、まだ検証段階だが、AIによるITシステム異常検知機能のサービス化に向けた取り組みが行われている¹¹。

故障の原因特定、及び、回復措置実施では、検知した異常に対し、対応方法が定型化されている場合、その手順に従い実施する。一方、定型化されていない場合、熟練者の暗黙知や過去の類似故障に関する対応履歴が頼りとなる。また、定型化自体にも稼働を要するという課題がある。そこで、AIを活用することで、故障時のシステムやネットワークから得られる情報から過去に起きた類似故障情報の自動提示や¹²、事前に故障時のネットワーク装置等から発せられるアラーム等のイベント情報と故障要因との関係を学習し、定型化作業を自動化することで¹³、故障対応者を支援する取り組みが行われている。

このように、故障対応へのAI適用は、自社内での故障対応業務の稼働軽減、効率化が主目的となっている。今後は、将来の保守人員減への対応としての位置づけもあるが、ネットワーク上で提供するサービスも多様化が進み、保守者が対応すべき内容は、ますます複雑化しているため、AI活用による保守者支援が進んでいくと考えられる。

- 2) 通信リソースの制御・オーケストレーションの自動化

⁸ https://www.hokudai.ac.jp/news/pdf/210204_pr_re.pdf

⁹ https://www.nttdocomo.co.jp/binary/pdf/corporate/technology/rd/technical_journal/bn/vol27_1/vol27_1_006jp.pdf

¹⁰ <https://optage.co.jp/recruit/mid-career/cross-talk/technology.html>

¹¹ https://www.hinemos.info/news/AIOps_202010

¹² <https://xtech.nikkei.com/it/atcl/news/17/101302437/>

¹³ <https://www.ntt.co.jp/news2017/1701/170130a.html>

エンタープライズユーザ等の通信ネットワークに対するニーズに対し、ユーザが直接、操作や変更ができるインタフェースをネットワークで提供しているケースでは、ユーザにより、オンデマンド、もしくは、早期にニーズを満たすことが可能となる。しかし、通常、ユーザが操作できる範囲は限定的であり、それ以外のケースでは、ネットワークやシステムのことが分かる専門家が要件を整理し、ネットワークや機器の設定に反映する等、人手を介するため、即座の対応は難しいという課題がある。そこで、AIによりサービス要件の分析を行い、必要なネットワークリソースの情報に自動変換することで、人手を介さずに、算出されたネットワークリソースの最適配分を行うネットワーク自動最適制御技術の研究開発が進んでいる¹⁴。

また、電話のような音声サービスであれば、会話が成立するために、遅延やゆらぎで満たすべき品質条件がある。一方で、データ伝送では、スループットが重要というように、サービスにより、通信ネットワークに求める品質条件は異なる。最近では、対戦型のオンラインゲーム等、より低遅延性を求めるサービスや、IoTのように無数のセンサとの同時接続が必要なサービスもあり、サービスが求める品質条件は、ますます多様化している。

現在、提供されているNGNで見ると、4つの転送品質クラスが利用可能であるが、音声サービスや映像配信サービスを念頭に、優先制御するための転送品質クラスを定義しており、今後、サービスにより求める品質条件がますます多様化していくと、新たなサービスが登場する都度、品質条件を設定し、多様なサービスの品質条件を満たすリソース割り当てを適切に行うことは困難になっていくと想定される。

そこで、AIを活用してネットワークを流れるトラフィックの特徴から様々なIoTサービスを自動的に分類することで、サービス特性に適した帯域等のネットワークリソース割り当てを可能とするソリューションが提供されている¹⁵。物理的には一つのネットワーク上で、ネットワークスライシング技術により品質条件の異なる複数のサービスを提供可能となりつつあるが、各ネットワークスライスに、どのようにリソースを割り当てるかは、割り当てルールをあらかじめ決めておく必要があり、品質条件が異なるサービスの増加に応じ、スライス数が増えると、リソースの最適な管理が困難になることも想定される。楽天モバイルでは、ネットワークスライシングの利用状況のデータに対し、AIを用いて予測、変化があった場合は主原因特定、リソースの最適化を行うことで、利用者の要望に対し、短時間できめ細かく応えられるシステム環境の実現を目指している¹⁶。ユーザのニーズにオンデマンドで対応していく上で、AI導入がさらに進んでいく領域と考えられる。

3.2 電力設備における適用事例

電力は、社会活動を支える基幹インフラであり、停電となった際の影響は非常に大きく、生活を含め、社会に混乱をまねくことになる。

近年、SDGsやESGに代表される世界的なカーボンフリー化を目指す機運の高まりもあり、各種再生可能エネルギーの適用拡大と共に、国内における発/送電分離や電力自由化等、電力を取り巻く環境や制度が大きく変化しており、電力の安定供給を継続する上で留意すべき事項が増加傾向にある。

電力分野におけるAI適用事例として、

- 1) 電力需給予測
- 2) 設備の予知保全、劣化推定
- 3) 設備運用の効率化

を取り上げた。

- 1) 電力需給予測

¹⁴ https://www.soumu.go.jp/main_content/000538848.pdf

¹⁵ https://jpn.nec.com/press/201802/20180207_02.html

¹⁶ https://www.soumu.go.jp/main_content/000696885.pdf

電力は、常に需要と供給のバランス（同時同量）が取れていることが要求され、需給バランスが崩れた場合、電気の品質である周波数が乱れることになる。供給が需要を上回ると周波数が上がり、需要が供給を上回ると周波数が下がるため、周波数の安定度が確保されない場合、安全装置が発動し、発電が停止、停電に至る場合がある。地震が起因であったが、2018年9月に北海道で発生した大規模停電（ブラックアウト）も周波数変動から大規模停電に至った例である。

電力需要は、季節、気象状況や時間帯により変動し、電力供給は、発電所や送配電設備のトラブル、太陽光や風力発電は天候条件等により変動する。そこで、電力会社は、予め作成した発電計画をベースに、需要変動に応じて発電量を変え、需要と供給を一致させる営みを継続的に行っている。

電源リソースが多様化・分散化する中で、供給側の制御対象が増えていることもあり、一層の需給予測精度の向上が求められており、AIを適用する例が散見される。

東芝は、研究開発中ではあるが、2012年から2014年の3年間の東京電力管内の気象データと電力需要データを学習データとしてAIに学習させ、2015年1年間の電力需要を予測した結果、従来の予測と比較して予測誤差の改善を報告している¹⁷。

多量の気象データをもつ気象予測会社は、電力需要と気象状況との相関関係に着目し、主に新電力事業者向けに、AIを適用した再生エネルギー発電量予測や地域毎の需要予測を行い、電力需給予測サービスを提供している^{18,19}。

2) 電力設備の予知保全、劣化推定

電力設備の予知保全、劣化推定にAIを適用する検討も行われている。

変電設備の画像・映像データや、機器稼働音を学習することで、通常とは異なる油漏れ状況や、振動や動作音を検出し、装置状況、ベアリング等の損傷や劣化を判別・検出している²⁰。

従来、変電設備は定期巡視により監視を行っているが、AIの適用により、巡視時間の大幅な削減が可能になるとしている。

また、太陽光発電の場合、ストリング電力値（太陽光パネル同士を接続した回路ごとに発電した電力値）をAIにより判定することで、発電低下等の異常を通知するシステムが開発されている²¹。従来の監視システムでは、発電量低下を一定の閾値で判定していたが、発電量は周囲環境の要因によっても変動することから、判定の信用性に課題があったが。このため、AIの適用に際して、他の要素も加味した判定を可能とし、これによって判定の信頼性向上が可能となっている。

風力発電においても、設備の損傷進展実験から学習データを取得し、AIに学習させることで、異常兆候の検知が可能になるとの報告がある²²。風力発電の風車は、ダメージを受けた後に交換する場合、交換計画の立案から交換部品の手配や施工とその後の正常性検証等、長期間の風車停止（発電停止時間）が必要となるが、予兆監視を行うことで、計画的な交換作業が可能となり、発電停止時間の大幅な短縮が図れ、結果として設備利用効率の向上が確認されている。

火力発電所では、大型ガスタービンに設置された複数センサからのデータを使用して、データ間の相関関係をMT法による検知とデータ間のパターンマッチングにより、故障箇所推定の可能性が報告されている²³。

¹⁷ https://www.toshiba.co.jp/tech/review/2019/05/74_05pdf/a06.pdf

¹⁸ <https://www.jwa.or.jp/service/weather-and-data/weather-and-data-02/>

¹⁹ <https://jp.weathernews.com/news/31676/>

²⁰ <https://www.nttdata.com/jp/ja/news/release/2018/121701/>

²¹ <https://sei.co.jp/company/press/2018/12/prs108.html>

²² https://www.nedo.go.jp/news/press/AA5_100939.html

²³ https://power.mhi.com/jp/randd/technical-review/pdf/index_54j.pdf

また、予知保全において、複数センサからのセンシングデータの関係性を学習し、異常検知を行うインバリアント解析技術に関する適用報告もある²⁴。

3) 設備運用の効率化

設備の運転管理や運転方法にAIを適用することで、設備運用の効率化検討も行われている。

従来、火力発電は、現場のオペレータの経験を基に、運転管理・事業運営を行っているが、AIの適用により、より高効率な管理・運営が出来るものと期待されている²⁵。水力発電においても、IoTやAIを活用することで、運用の効率化や、天候情報を併せて解析することによる発電電力量の増加検討も行われている²⁶。

3.3 モビリティサービスにおける適用事例

MaaS (Mobility as a Service) に関する AI 活用事例として下記がある。

AI 活用によって実現可能な技術の代表例として、自動運転が挙げられる。この自動運転システムでは、道路標識や歩行者、対向車といった自動車の周辺状況を把握する認知プロセスにおいて AI が多く利用されている²⁷。また、周辺状況を画像として取込み、大量のデータから重要な特徴量を抽出し、認識した対象が「人」か「障害物」を分析している。障害物を検知した時、急ブレーキを踏む等の判断や、事故リスクや危険可能性の算出する予測技術「プレディクション (prediction)」において AI が活用される事例もある²⁸。このように、自動運転では、AI が正確な物体検知を常にリアルタイムで行う必要があり、AI 技術の精度と、大量のビッグデータ解析処理を行う高速な計算資源が求められている。

また、自動運転車は、走るデータ収集装置といわれ、カメラ、ミリ波レーダー、赤外線レーザースキャナ (LiDAR) 等、数多くのセンサを搭載し、車内カメラやマイク、脈拍センサ等で、ドライバー生体情報から、人の行動や、感情、健康状態等も、今後、推定可能とされる。

以下に3つの自動運転技術の実用例について示す。

・テスラの先進運転システム「オートパイロット (AutoPilot)」

テスラの自動運転システムは、Lidar (光検出と測距) や詳細マップは不要で、カメラ画像のAI解析だけで自動運転を実現している。また、Teslaは無人ライドシェア事業も展開している。2019年4月、完全自動車運転の実現に向けて新機能「オートパイロット・レーンチェンジング」を投入し、この機能は、クルマが高速道路の各レーンの状況をセンサで把握し、自動的に車線を変更するものとなっている²⁹。また、信号や一時停止標識で自動ブレーキをかけるように進化している。

・MobileEye (モービルアイ)

MobileEyeは、インテルの子会社で、単眼カメラを用いて衝突事故防止・軽減を実現し、先進運転支援システム(ADAS)の発展に貢献するテクノロジーカンパニーである。人工視覚イメージ処理技術、マルチテクノロジーアプリケーション、情報技術を組み合わせ安全運転実現のために幅広い解決策を提供している。CES 2020では、カメラだけを車両に設置して自動運転を実現する動画を公開した。³⁰

²⁴ <https://jpn.nec.com/techrep/journal/g15/n02/pdf/150213.pdf>

²⁵ <https://www.nedo.go.jp/content/100896705.pdf>

²⁶ https://www.kepco.co.jp/corporate/pr/2018/0918_2j.html

²⁷ <https://nissnad-digitalhub.com/articles/ai-for-self-driving-structure/>

²⁸ https://jidouten-lab.com/y_1298

²⁹ https://jidouten-lab.com/u_robot-taxi-matome

³⁰ <https://newsroom.intel.co.jp/news-releases/ces-2021-mobileeye-avs-on-move/#gs.277wd5>

・本田技研工業：高精度3次元地図データ活用し、レベル3の自動運転車両を発売
日本において自動化レベル3（条件付き自動運転）自動運転システム「Honda SENSING Elite」を搭載した新型「レジェンド」を3月に発売している。このシステムは、ゼンリンの「高精度3次元地図データ(HDマップ)」を採用している。³¹

4 各所団体における検討状況

4.1 通信設備関連団体における検討状況

TMForumのAutonomous Networks Project³²では、自律的なネットワークのロードマップやclosed-loopのアーキテクチャ、ケーススタディの議論が進んでいる。また、TMForumのAI Management Standards projectの中で、企業がAIの複雑さを管理し、継続的に説明責任を果たし、安全かつ責任を持ってAIを使用するのを支援するAIチェックリストが制定されている³³。これは、運用しているドメインの複雑さを管理する方法として、航空業界でパイロットが安全確保のために標準化されたチェックリストが利用されていること、あるいは外科医の手術において安全性を高めるチェックリストが利用されていることを参考としている。

AIチェックリストは、AIのライフサイクルの6つの段階(調達、開発前、開発後、配備、利用時、利用終了時)ごとに目的や各段階で確認すべき項目を示しており、継続的に検証および更新される「動的」ドキュメントで、AI導入から終了まで、考慮すべき項目をチェックリストで示している。例えば、AIシステムに関するテストを計画通り実施し、結果を保持しておくことは、学習の透明性やAIが対応できる範囲／対応できない範囲の根拠を作ることと言える。また、運用での定期的なレビューは、AIシステムが想定通りの品質を確保できていること、及び、インシデントが管理できていることを示すことにより、AIシステムをコントロール下に置いて運用していることを示すことと言える。

現在、通信キャリアでのAI導入は、キャリア内での業務効率化を主目的にしているケースが多く、AIシステム主導で、お客様へのサービス提供を実施する段階には、まだない。今後、サービス多様化により、オペレーションの複雑化が想定される中で、先のAI適用事例であげた、故障対応の自動化や通信リソースの制御・オーケストレーションの自動化を進めていくことは、キャリアにとっても重要な課題である。チェックリストの活用により、AIシステム導入準備や運用時のトラブル対応、改善活動を円滑に進め、導入や運用で得た経験を記録し、次のAIシステム導入に生かすサイクルを確立することは、AIシステムを導入し、活用していく土壌を作っていく上で意味がある取り組みと言える。AIシステム本格利用期に向け、AIシステムの挙動がユーザの効用に与える影響を評価し、AIシステムを持続的に発展可能とするチェックリストに作り上げていくことが期待される。

ITU-TでもAutonomous Networks に関するFocus Groupが立ち上がり、今後、議論の活性化が期待される³⁴。ETI-ZSM でも、ゼロタッチでのネットワークやサービスマネジメントのアーキテクチャの議論が進んでいる。このように、通信設備関連の各種団体で、AIを活用した自律的なネットワークや運用自動化に関する議論がされており、今後、AI活用が進む中で、信頼性や説明責任に関する議論も進んでいくと思われる。

4.2 電力設備関連団体における検討状況

³¹ <https://car.watch.impress.co.jp/docs/news/1309878.html>

³² <https://www.tmforum.org/collaboration/autonomous-networks-project/>

³³ <https://inform.tmforum.org/insights/2020/06/new-ai-checklists-address-complexity-and-accountability/>

³⁴ <https://www.itu.int/en/ITU-T/focusgroups/an/Pages/default.aspx>

電力インフラに関する省令の「電気設備に関する技術基準を定める省令第四十六条」において、火力発電所は技術員による常時監視が義務付けられている。また、送電設備も定期的な点検や監視を技術員により実施されることが求められている。

電力分野においても、熟練技術者の高齢化や人員不足等が懸念されており、発電所や送電設備の点検や監視業務を今の水準で行うことが困難になりつつあり、発電所や送電設備の保守・点検業務へのAI適用が検討されている。しかしながら、現状では、AI適用の検討や有効性確認の段階にあり、広く運用に供される段階には至っていないこともあり、電力装置の予知保全、推定劣化にAIを適用した際のAIに対する信頼性、公平性に対する標準的な考え方の一般化は行われていない。

他方、政府が中心となり電力インフラに対するスマート化推進の検討が行われており、その中で省令改定も含めた検討が行われており、近い将来、電力インフラに対する標準化指針等が出されると考えられる。

4.3 モビリティサービス関連団体における検討状況

自動運転やAI、オープンデータ等を活用し、従来型の交通・移動手段にシェアリングサービスも統合した次世代のモビリティサービスが、グローバルに普及しつつある。

その中心となる自動運転技術については、地理データ、移動データ、通信データなどの異なるデータを繋げるアーキテクチャ、相互運用性を考えていく必要があり、サービスやデータ品質・プロセス等の標準化に関する議論が開始されている。

ITU-T FG-AI4AD (Focus Group on AI for autonomous and assisted driving) では、自動運転用AIを検討するITU-Tのプレ標準化グループが2020年1月発足し、道路上のAIの評価に関する仕様やガイドライン、ギャップ分析、アーキテクチャや、SDGs達成に向けたロードマップ等の策定が開始されている。

ITU-T FG-AI4ADが策定中のドキュメントを以下に示す。

- ① TR01: “Automated driving safety data protocol – Specification”
- ② TR02: "Automated driving safety data protocol – Public safety benefits of continual monitoring”
- ③ TR03: "Automated driving safety data protocol – Practical demonstrators”

(出典) ITU-T FG-AI4AD <https://extranet.itu.int/sites/itu-t/focusgroups/ai4ad/input/FGAI4AD-I-102.pdf>

TR01(Automated driving safety data protocol – Specification)では、自動運転のAIの振る舞いに関連した通信、プロトコル仕様について検討が行われている。特に、自動運転では、カメラ、ミリ波レーダー、赤外線レーザーキャナ (LiDAR) 等、数多くのセンサから発生するデータをリアルタイムにモニタリング・分析し、その情報を用いて意思決定できること、自動運転に関わるシステムはハードウェアとソフトウェアは独立し、AIを用いる場合は説明可能な最低限の基準が必要不可欠であると述べている。

また、ユーザに対して自動運転システムの安全性を提供するにあたり、ユーザに対する透明性及び説明責任と、独自の知的財産権の保護のバランスが重要なファクタと述べている。現在、TR01では、自動運転のためのAIの安全性と倫理に関する“モリー問題 (The Molly Problem)” (自動運転車が事故を起こし、目撃者がいない場合、衝突時の自動運転 (AD) ソフトウェアの挙動に関する課題を提起) を中心に議論が行われ、そこから自動運転の振る舞い評価の要求条件として、“衝突を検知するソフトウェア”、“衝突時緊急停止可能なソフトウェア”等が議論されている。しかしながら、車両データと自動運転のソフトウェアのデータに関してはサイロ化され、安全検知できるようなデータ統合までは実現できていないのが現状である。

TR02("Automated driving safety data protocol – Public safety benefits of continual monitoring)では、The Data Ethics Commission (DEK)の損害の重要度 (Severity of harm) について触れ、信頼性・安全性の確保のため、開発者や製造業者が、自動運転システムやAIなどの関連技術を一貫して評価できる尺度が必要であるとされてい

る。この重要度は害を及ぼす可能性をLevel1からLevel5の5段階で表現し、自動運転やAIで活用するデータ（個人データなどの法的に保護されるデータ）の扱い、データの品質、個人への影響やその影響した人数、社会への影響など、重要度の評価への考慮について議論が進められている。

また、TR02では、法的な観点についても触れており、特に、the EU Product Liability Directive (PLD)に関する見解について分析している。自動運転も含めた事故の責任は国によって異なる点（国ごとの交通ルールに従う）や、自動運転車の製造業者はソフトウェアや部品、また、それらを組み立てて市場に投入するプロバイダと多岐にわたるため、責任の所在が複雑になるとともに、事故が発生した際には、被害者がその原因を調査する証明の負担も大きいという課題を述べており、自動運転の信頼性の観点で課題があるとしている。

さらに、ソフトウェア（AIを含む）は、自動車の一部なのか不明確であるとともに（事故が起こった際には裁判所の判断となる）、ソフトウェアのアップデートに関しても、更新する行為が新たなプロダクトとして認識すべきなのかなど、法律やガイドライン策定に向けても今後議論が必要な点と述べている。一方で、自動運転の技術に関しては、まだ様々な課題があるものの、その課題によって技術革新が阻害されてはいけないため、国家的に開発リスクの保護が受けられる施策も、将来的な自動運転の品質や信頼性のためには必要であると言及している。

TR03(Automated driving safety data protocol – Practical demonstrators)ではOpenPilotとAutowareの2つのオープンソース自動運転スタックのシミュレーションベースのデモンストレーションの概要を記載している。Autowareの開発者であるTier IVは、計画、認識、および制御インタフェースの緊密な結合を減らすために、AVソフトウェアスタックの新しいアーキテクチャ設計を提案し、センシング、知覚、ローカリゼーション、計画、制御の6つのモジュールを定義している。自動運転車のテクノロジーの効率的なリスク分析のシミュレーション検討として、Trustworthy Searchツールを用いて、ブラックボックス方式でシステムと対話し、テスト対象のソフトウェアの整合性と機密性を維持するテストも検討している。

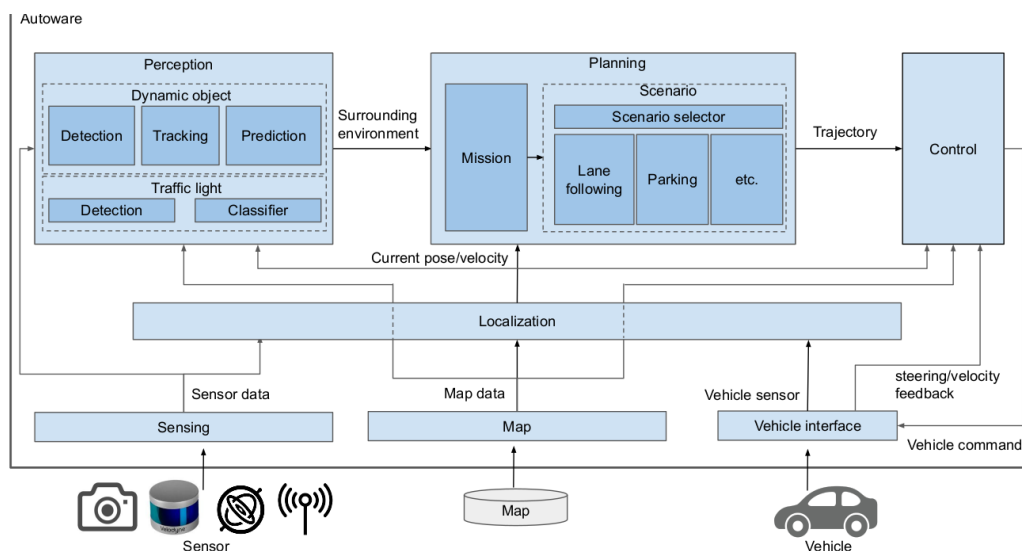


図2 Automated driving safety data protocol – Practical demonstrators (出典) ITU-T FG AI4AD TR03

また、European Commissionにおいては、“Ethics of Connected and Automated Vehicles”と題して、自動運転に関する安全性、プライバシー、公平性、説明責任について、20個のリコメンデーションをまとめている。自動運転システムに関しては、自動運転が必ずしも安全ではないという考えに基づいた制御機構（人でも制御できる等）の必要性、情報のプライバシー保護（GDPRの個人データの取扱い）の観点では個人データに関するユーザ合意の必要性、また、AIを活用したデータ分析では個人特定が可能になるケースもあることからアルゴリズムに対して、透明性の確立に向けた業界標準・ベストプラクティスの策定の必要性があると述べている。

上記の分析による個人特定に関しては、新たなプライバシーリスクとなりうることから、AI分析による新たな個人保護のガイドライン化も必要となる。また、AIを活用したデータ分析においては、アルゴリズムのブラックボックス化やデータ偏りによる差別化・不平等な結果（例：優先制御や不平等な価格設定など）を生じる可能性もあり、アルゴリズム設計や運用に対して、監視・評価する手段も今後の課題と定義されている。具体的には、アルゴリズムの透明性を担保する仕組みとして、監視ツールの開発やユーザからのフィードバック機能の充実（例：政策的な援助も含め）や、AIに関わる説明責任機能は必要不可欠であるとされている。合わせて、自動運転システムの適切な開発や利用促進のために、開発者やユーザの道徳的知識も必要不可欠であるとし、スキル向上のトレーニングや公的資格などの仕組み作りに関して提言している。

自動運転の安全性に関する他の標準化動向として、自動車向けの機能安全規格であるISO26262と、SOTIF（Safety Of The Intended Functionality）がある。ISO26262ではディプロイメント（配置展開）前の機能の安全性を、ISO SOTIF³⁵（SAFETY OF THE INTENDED FUNCTIONALITY）ではディプロイメント前後の特定機能における安全性（例えば性能限界や誤使用・誤操作等）が議論されている。

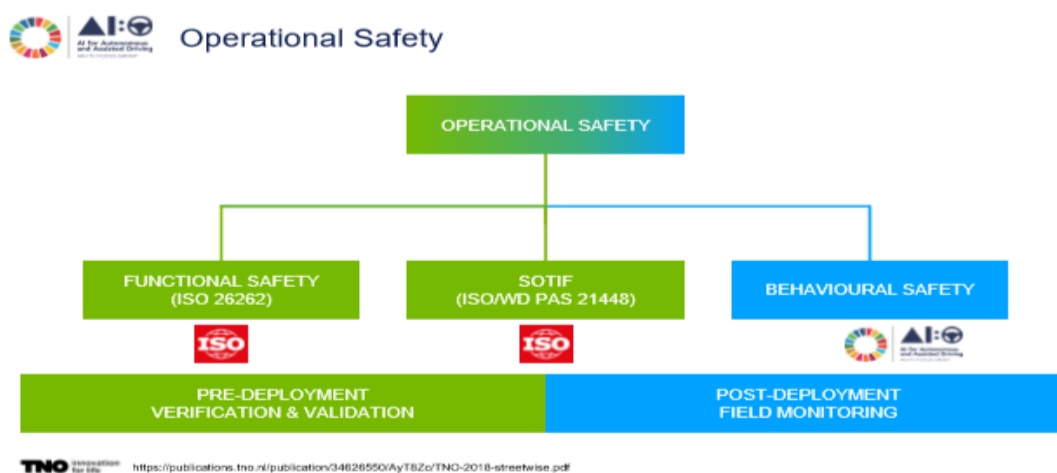


図3 Operational Safety （出典）ITU-T Update on FG-AI4AD activities³⁶

5 GAP分析

5.1 通信設備関連団体の状況とのGAP分析

AIの活用検討は通信キャリア各社で行われているが、キャリアごとに利用している通信設備の機器は異なり、また、ネットワークの構成も異なることに加え、ネットワーク構成や運用方法自体が機密情報やノウハウを含み公開が難しいため、データの共通化が難しく、AIシステムの適正学習や信頼性に関する統一的な議論は進んでいない。さらに、通信設備では、増減設や通信装置のファームウェアのアップデートが頻繁に発生するため、AIモデルの安定的な運用が難しいという課題もある。このように、キャリア内での業務支援を目的としたAI活用でもまだ課題は多いが、将来、AIによるサービス提供にシフトしていくと、故障対応、及び、サービス受付やプロビジョニングの領域に関しても、例えば、以下の場面では、AIが実施した内容に関し、ユーザへの説明責任が求められる。

- SLAを満たした運用がなされていること
- サービス影響を未然に防ぐための点検業務がされていること
- 故障発生や影響範囲が即座に通知され、回復見込み時間、及び暫定措置が提示できること

³⁵ MONOist （参考） <https://monoist.atmarkit.co.jp/mn/articles/1708/02/news030.html>

³⁶ <https://extranet.itu.int/sites/itu-t/focusgroups/ai4ad/input/FGAI4AD-I-102.pdf>

- (長時間故障の場合) 故障対応に時間がかかった理由
- (故障回復が段階的に進む場合) 故障回復の順序
- 優先制御での公平性、他ユーザの制御により受けた影響
- 輻輳制御での制御対象とする通信の選定、及び、制御の理由
- 品質や料金の観点から適正にプロビジョニングが行われていること
- ユーザの要求がサービス受付に正しく反映されていること

人では対応しきれない膨大なデータをAIに活用することで、通信分野でも、人では気づけなかった異常の発見や、同時に対応しきれなかったことに対応できる可能性があり、AIの強みを活かすことで、今よりも、よりよい通信サービスが提供できていると説明できる日がいずれ来ると想定している。

5.2 電力設備関連団体の状況との GAP 分析

電力分野の運用・管理へのAI適用の検討が行われているが、AI適用に際して品質/信頼性に関する取組は少ないのが現状である。

そこで、今後の電力分野へのAI適用にあたって、品質、信頼性の観点で考慮すべき点の整理を行った。

先に示した3種のケースで共通するものとして、「使用するパラメータの種類」、「適正な学習データの使用」、「使用するモデルの品質評価」等、一般的なAIの品質を評価するための項目に加え、電力需給予測の場合では、「誤差の評価方法とその値」、設備の予知保全、劣化推定の場合では「結果に対する信頼度」、「結果に至った理由」、設備の運用効率化の場合では「効率化の度合い」、「効率化が可能な理由」等が挙げられる。

なお、電力設備が対象ではないが石油プラントの保安にAIを使う場合の評価について、「プラント保安分野AI信頼性評価ガイドライン」³⁷が公開されており、電力インフラへの応用が考えられる。

本ガイドラインでは、AIの信頼性評価として、「利用時品質」、「外部品質」、「内部品質」に分類されており、品質確保手順としては、下記の3ステップが示されている。

- ① AIで実現したいこと(利用時品質)の設定
- ② 利用時品質を満たすために機械学習要素に求める性能(外部品質)の明確化と要求レベルの設定
- ③ レベル毎の要求事項に基づいた機械学習要素(内部品質)の作り込み

この階層的な信頼性評価方法について、具体的なユースケース(設備劣化診断、異常予兆検知等)が詳述されている。

しかし、本ガイドラインには、定量的な指針が記載されておらず、今後の品質評価尺度の一般化には定量的な指針が必要と思われるが、各所で劣化診断、予知検知の研究や開発が行われていることから、十分なデータ蓄積と、今後のAI適用検討の進化により定量的な指針が設定されると思われる。

先に述べたインバリエント解析技術³⁸は、各種センサの時系列データを見える化をすることで、理由や根拠の可視化が可能なホワイトボックス型AIでもあり、本形式によるAIの信頼性を満たす事も考えられる。

5.3 モビリティサービス関連団体の状況との GAP 分析

4.3章で述べたように、AIを画像認識等に適用し、自動運転の実現に向けた技術開発が進んでいる。しかしながら一方で、カメラで撮影した周囲の映像から画像認識を行って自動運転を実行するテスラの自動運転システムに対して、一般的な道路標識に変更を加えるだけで、自動運転車のシステムが混乱をきたし、標識の意味を誤解釈する可能性があるという報告も上がっている。[5]

³⁷ <https://www.meti.go.jp/press/2020/11/20201117001/20201117001.html>

³⁸ <https://jpn.nec.com/ai/analyze/invariant.html>

米電気自動車（EV）大手テスラの自動車は、2件の死亡事故を起こしている。いずれも同社の自動運転機能「オートパイロット」使用中に発生しており、自動運転産業でいわゆる「エッジ・ケース」として知られる例、つまり自律制御ソフトウェアが人の認識とは異なる判断、誤検知をしたために、車両が予測不能な動きをする恐れがある。

上記の事故は、自動運転に関わる法的ルールの整備及び一貫した明確なガイドライン化が進んでいない状況で起こった事象である。この課題を解決するために、4.3章で述べたような、標準化活動（ITU-TFG-AI4ADやEuropean Commission等）では、自動運転（AIを含む）の品質・信頼性のためのガイドライン策定や、事故が発生した際の責任にかかわる今後の課題に関して議論が行われている。

一方で、更なる信頼性の向上には、ガイドライン化や法的ルールの整備だけではなく、実際に起こった事象やユーザからのフィードバックに基づいて、技術的改善、ガイドライン及び法的ルールの改善が継続的に実行できる仕組み作りが必要となると考える。また、自動運転では開発者・製造業者（ソフトウェアとハードウェア）、利用者など利害関係者が多岐にわたるため、自動運転の品質/信頼性、責任の所在の考え方等に対して、一貫した理解を持つことも、将来的に品質や信頼性の向上に大きく寄与する仕組みであると考えられる。

6 まとめ

スマートフォンやDigital Transformationの普及により、日々の生活に必要な様々なデジタルサービスだけでなく、ライフライン等インフラ設備への、AIの活用が検討され始めている。2020年にサービスが開始された5G（5th Generation mobile network）の普及により、今後、通信インフラとサービスの融合が一層加速化し、データを活用した新たなサービス創出が想定されている。

このデータ利活用におけるキーテクノロジーがAIであるが、AI適用の普及には、AI自体の品質、信頼性と説明責任への対応が不可欠であると共に、その在り方が問われている。

本TRでは、国民生活に欠かせないインフラ（通信と電力）とMaaS（自動運転）に焦点を当て、AIの活用事例と、品質・信頼性・説明責任に関わる技術的、標準的な動向を明らかにするとともに、今後のGAPと展望について記述した。

3事例とも、AIの利活用検討は進められているが、AIの品質、信頼性につながる標準化やガイドライン及び法律の整備は緒に就いたばかりであることが、今回確認できた。

しかしながら、AI普及の加速には、各ステークホルダがAIに対する目的やその適用や取組など、領域別の価値観を統一していくことも重要であると共に、今後のAIの品質、信頼性と説明責任に対して大きな貢献をすることになると考えている。

今後、今回の調査領域における取組の進展具合や調査対象領域を拡大による改版を視野に入れたいと考えている。

付録 FG-AI4AD モリ－問題の調査の結果（2020年10月11-20日）

No.1	Would you expect...	Yes	Unsure	No
1	the software <u>to be aware</u> of the collision 衝突を認識するソフトウェア	97%	2%	1%
2	the software <u>to stop</u> at the collision site 衝突現場で停止するソフトウェア	94%	4%	2%
3	the software <u>to indicate a hazard</u> to other road users 他の道路利用者に危険を示すソフトウェア	97%	2%	1%
4	the software <u>to alert emergency services</u> 緊急サービスに警告するソフトウェア	94%	5%	1%
5	the software to recall the <u>time</u> of the collision 衝突の時間を記録するソフトウェア	99%	n/a	1%
6	the software to recall the <u>location</u> of the collision 衝突の場所を記録するソフトウェア	99%	n/a	1%
7	the software to recall <u>when</u> the collision risk was identified 衝突の危険性が特定された時を記憶するソフトウェア	93%	6%	1%
8	the software to recall <u>if</u> Molly was detected モリ－が検出されたかどうか、記憶するソフトウェア	96%	3%	1%
9	the software to recall <u>when</u> Molly was detected モリ－が検出された時を記憶するソフトウェア	96%	2%	2%
10	the software to recall <u>if</u> Molly was detected as a <u>human</u> モリ－が人間として検出されたかどうかを記憶するソフトウェア	91%	6%	3%
11	the software to recall <u>when</u> Molly was detected as a <u>human</u> モリ－が人間として検出された時を記憶するソフトウェア	90%	7%	3%
12	the software to recall <u>whether</u> mitigating action was taken 緩和措置が取られたかどうかを記憶するソフトウェア	98%	1%	1%
13	the software to recall <u>when</u> mitigating action was taken 緩和措置が取られた時を記憶するソフトウェア	97%	2%	1%
14	the software to recall <u>what</u> mitigating action was taken どのような緩和措置が取られたかを記憶するソフトウェア	96%	3%	1%
15	similar recall abilities for <u>near-miss events</u> ニアミスイベントに対する同様の記録能力	88%	5%	7%
16	expect <u>driving</u> to be <u>prohibited</u> for software <u>without recall capability</u> 記録機能のないソフトウェアでは運転が禁止されることを期待する	72%	15%	12%