

## AIの標準化概観 —ELSI対応を中心に—

日本電気株式会社 標準化推進部 えがわ たかし  
江川 尚志



### 1. はじめに

AIは、一般社会では2014年前後から始まった「こんなことができるようになった」「人を越えた」といった驚きに彩られたフェーズを過ぎ、具体的な応用を案出し、実用化する上での問題点を一つひとつ解決するフェーズに移りつつある。

この数年間、AIは社会を大きく変える技術であるとの認識の下、倫理や法律、経済や教育を含む様々な社会的影響が各国で分析され、それに基づく社会合意レベルの大きな枠組みが設定されつつある。またその大きな枠組みを実務レベルに落とす作業の一つとして、2017年にはAIを扱うグループの設立が標準化団体で相次いだ。

AIの標準には、ICTシステムの標準化で常に必要とされてきた用語や品質尺度といった従来標準の延長線上にある標準と、倫理や信頼 (trustworthiness) といったAIならではの標準とがある。

前者のICTシステムとしての標準化としては、[米国R&D戦略]が7個の戦略の6番目として「基準及びベンチマークを通じてAI技術を計測し評価」を挙げ、従来の観点から必要になる様々な標準を列挙している。

後者のAIならではの標準を象徴するのは「倫理」[trustworthiness][ELSI]といった言葉である。AIに対する高い期待の反動として「AIは余りに高性能であり何をするか分からない、倫理に反することをしたり社会を破壊したりするかも」という不安が人々の間に生まれたと筆者は感じている。これはICT関係者に「せっかくの技術やサービスを人々に使ってもらえないかもしれない」との懸念を引き起こし、使ってもらったため人々のAI技術に対するtrustworthinessを高める標準を、との議論につながっている。例えば「事故が起きた時、原因を究明可能とするために記録すべき情報を定める(透明性)標準」の議論である。これは一般には遺伝子組み換え技術の登場に伴って生み出されたELSI (Ethical, Legal, Social Issues)と呼ばれる分野であり、AIの標準化を特徴付けている。

本稿では後者、倫理、trustworthiness、ELSIといった話題に関連する標準を中心にAIの標準化を概観する。

### 2. AIで必要となる標準化

AIの実用化に際しては様々な側面での課題が生じ得る。

内閣府人工知能と人間社会に関する懇談会では、移動、製造、個人向けサービス (医療と金融を含む)、対話・交流 (コミュニケーション) の各分野について倫理、法、経済、社会、教育、研究開発の各側面から生じ得る論点を整理している [懇談会]。それによれば、例えば、移動 (自動走行車) では倫理的論点として危機回避における優先順位を誰がどう決めるか、法的論点として自動運転中に起きた事故の責任や、自動運転車に搭載されるだろう高性能カメラによるプライバシーの侵害、経済的論点として保険の在り方、等々多数の問題を挙げ、分野にまたがる普遍的な論点へと整理している。

これら様々な分野に深く影響する課題を処理するには、大きな枠組みとして社会の基本合意レベルの大原則を定めると同時に、実務者の日々の指針となり得るレベルで細部を決定していく必要がある。

大きな枠組みを設定する国家レベルの議論で先行したのは日本で、総務省AIネットワーク化推進会議は2017年7月、AI開発ガイドライン [総務省ガイドライン] を策定して国際的な制度の協調を図る活動を開始した。本ガイドラインは9個の原則を3つに分けて述べている。

(主にAIネットワーク化の健全な進展及びAIシステムの便益の増進に関する原則)

- ・連携の原則；開発者は、AIシステムの相互接続性と相互運用性に留意する

(主にAIシステムのリスクの抑制に関する原則)

- ・透明性の原則；開発者は、AIシステム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
- ・制御可能性の原則；開発者は、AIシステムの制御可能性に留意する
- ・安全の原則；開発者は、AIシステムがアクチュエータ等を通じて利用者及び第三者の生命・身体・財産に危害を及ぼすことがないように配慮する
- ・セキュリティの原則；開発者は、AIシステムのセキュリティに留意する
- ・プライバシーの原則；開発者は、AIシステムにより利用者及び第三者のプライバシーが侵害されないよう配慮する
- ・倫理の原則；開発者は、AIシステムの開発において、人間の尊厳と個人の自律を尊重する

(主に利用者等の受容性の向上に関する原則)

- ・利用者支援の原則：開発者は、AIシステムが利用者を支援し、利用者を選択の機会を適切に提供することが可能となるよう配慮する
  - ・アカウントビリティ：開発者は、利用者を含むステークホルダーに対しアカウントビリティを果たすよう努める
- これに続いたのが欧州であり、欧州委員会の科学技術倫理グループは2018年3月、以下の9原則〔欧州9原則〕を発表した。
- ・ Human dignity (人間の尊厳)
  - ・ Autonomy (自律)
  - ・ Responsibility (責任)
  - ・ Justice, equity, and solidarity (正義、平等、連帯)
  - ・ Democracy (民主主義)
  - ・ Rule of law and accountability (法の支配とアカウントビリティ)
  - ・ Security, safety, bodily and mental integrity (セキュリティ、安全、健全な身体と精神)
  - ・ Data protection and privacy (データの保護とプライバシー)
  - ・ Sustainability (持続可能性)

この他にも民間レベルで様々な原則が発表されている。アシロマ会議は23か条からなる〔アシロマ原則〕を発表した。個社のレベルでもいくつか発表されている。

こうした枠組みは重要ではあるが、技術開発や商品販売の実務者にはより具体的な指針や規則が必要となる。民主主義の原則と言われても実務者は困惑する。Trustworthinessを技術開発の手順で担保するならば、守るべき手順を示したプロセス標準、商品販売に使うAIが倫理的観点から行っていない行為があるならばそれを示した標準やガイドラインが必要である。

以下本稿では、これら具体的な標準やガイドラインに向けた検討の現状を概観する。

### 3. IEEE GIEAIS

IEEEは最も早くからAIの倫理面の標準化に取り組んだ標準化団体の一つである。IEEE Standards Associationは2016年4月にThe IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (GIEAIS) (当初の名称はGlobal Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems) を立ち上げ、AIの倫理面に関する論点整理を開始した。議長は

Raja Chatila仏国立科学研究センター (CNRS) インテリジェントシステム&ロボット研究所 (ISIR) 所長、実務はIEEEコンサルタントで社会活動家のJohn Havens氏が仕切る。設立前から非公式に議論が行われており、設立時点で65名、現在は250名以上の法律、倫理、技術その他の専門家が主に英米から参加して討議している。日本からもアカデミアを中心に数名の参加者がいる。

その結果はEthically Aligned Design (EAD: 倫理的に整合した設計) という文書にまとめられるとともに、標準化で解決すべき課題が発見されたならばP70xxシリーズ標準化プロジェクトとして立ち上げ、標準化が進められている。

#### 3.1 EAD

IEEE GIEAISの主たる成果の一つであるEADは、第1版が2016年12月に、第2版が2017年12月に発行された。GIEAIS発足時点では7個の論点が設定されていたが、その後増え、EAD第2版発行時点では下記の13の分野が設定されている。

- ・ General Principles (一般原則)
- ・ Embedding Values into Autonomous Intelligent Systems (A/ISへの価値組み込み)
- ・ Methodologies to Guide Ethical Research and Design (倫理的研究と設計を導く方法論)
- ・ Safety and Beneficence of Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI) (汎用AIと超知性の安全性と受益者)
- ・ Personal Data and Individual Access Control (個人情報とアクセス制御)
- ・ Reframing Autonomous Weapons Systems (自律兵器の再定義)
- ・ Economics/Humanitarian Issues (経済及び人道的事項)
- ・ Law (法)
- ・ Affective Computing (感情を読み表すコンピュータ)
- ・ Policy (政策)
- ・ Classical Ethics in A/IS (A/ISにおける古典的倫理)
- ・ Mixed Reality in ICT (ICTでの複合現実)
- ・ Well-being (幸福)

様々な分野が挙げられているが、特徴的な活動として汎用人工知能と超知性の問題が挙げられていることが目を引く。汎用人工知能と超知性は特に欧米の参加者の関心が高い課題であるが、汎用人工知能はまだ開発途上の技術であり、具体的な性質や能力が不明確であることも手伝い、明確な議論になっているようには見えない。

EADは上記13分野の各々について10個前後の論点を挙



げている。例えば上記2番目の分野であるA/ISへの価値組み込みでは、EAD第2版は下記のように9個の論点を3つに分類して述べている。

#### A/ISのための規範の特定

- ・ 特定されるべき規範
- ・ 規範の更新
- ・ 規範の衝突と解決

#### A/ISへの規範の実装

- ・ 多数の規範の実装方式が現存し新方式も開発中
- ・ 透明性の実装と適用
- ・ 失敗は起きる

#### A/ISの実装の評価

- ・ 目標となるコミュニティの規範の人間やAIへの不平等な実装
- ・ A/ISが一部グループに対し不利益となるバイアスを持つ可能性
- ・ 第三者評価の問題

その上で各々の論点について背景と分析、勧告案、参考情報を挙げる。例えば最初の論点、特定されるべき規範は次のように分析される。

#### 背景と分析

- ・ 社会には法律に加え社会的、倫理的規範が存在し、それらは必ずしも文書化されておらず、個人レベルから国家レベルまで各々レベルごとに様々な規範がある
- ・ 特定のコミュニティにおいて広く観察される規範を特定することは可能と信じるが、全てを網羅した統一的な規範を作成するのは非現実的である
- ・ 規範を実装する際にはA/ISの用途ごとにコミュニティを明確にする必要がある。例えば自動運転とヘルスケアの規範は異なること、コミュニティ内部には下位のコミュニティ（一部の人間のみが参加しているコミュニティ）があり下位コミュニティの規範を尊重する必要があること、と同時に下位のコミュニティの規範は上位のコミュニティの規範には反してはならない
- ・ A/ISに適用される規範は必ずしも人間に適用される規範と同じではない

#### 勧告案

社会的、道徳的規範に従うA/ISの開発では、最初に行うべきはA/ISが適用されるコミュニティの規範、特にA/ISが行うタスクに関連する規範を特定することである。

以上を述べた上で参考文献を列挙する。

EADは現在第2版であるが、第3版に向けた改定作業が始まっており、現在は新規論点の募集が既存の論点のレビュー

と並行して行われている。第3版の完成は2019年の第2か第3四半期を目指している。第3版は最終版とし、そこでEAD作成という形での活動は終結する予定となっている。そのため第3版では、第2版で勧告案とされている記述は、勧告案ではなく勧告となる予定である。

#### 3.2 P70xx

EADで特定された論点には、教育で解決すべきもの、広報活動で解決すべきものなど様々な課題が混じっている。うち標準化で解決すべきとされたものはIEEEの標準化プロジェクトとして議論される。

次ページの表に示すように、現在、P7000からP7012まで13個の標準化プロジェクトが活動している。表に示した完成目標はProject Authorization Request (PAR) に記載された第1回スポンサー投票の時期である (IEEEでは技術的内容に責任を持つ親ソサエティのことをスポンサーと呼び、そこで投票され承認されると技術的には完成と見なされる)。全般に作業は遅れており、まだ完成した標準はない。

IEEEでは未完成の標準の内容はメンバー外には守秘義務があるため、以下はPARほかに基づく情報である。また将来展望も記述できないことをお許しいただきたい。

P7000 (倫理的設計プロセス) は、技術に関わる関係者がシステム開発を開始し、分析し、設計する時に倫理上の懸念を適切に扱う事ができるよう開発手順のモデルを定める。利用者としては、技術者に加え研究者、企業、大学、各種の組織など技術の具体化に関わる組織を広く想定する。

P7001 (透明性) は、透明性のレベルを測定し、試験可能にすることを目指す。具体的には設計者が「本製品の透明性はレベル3」などといった標準への適合レベルを評価可能とし、また設計者に対して透明性向上のための示唆 (センサや内部状態は安全なストレージに保存するなど) を与えることを目指す。

P7002 (データプライバシーのプロセス) は、従業員や顧客、その他の外部の関係者の個人情報を用いる製品やサービスでの、プライバシー上の懸念に配慮したシステム/ソフトウェアエンジニアリングプロセスの要求条件を規定し、関係者がプライバシー上のチェック、適切に行われている事の確認を可能とする。

P7003 (バイアス、差別) は、アルゴリズムの開発において「負のバイアス」(人種や性別など法的に禁じられている差別、法的ではない差別を共に想定) を特定し除去する方法論を規定する。このためにバイアスを制御するための検証用データセットの選択手順と基準、アルゴリズムの利用限界を規定し、伝え、対象外の利用を防ぐガイドライン等を記述する。

P7004 (子供と学生のデータのガバナンス) は、AIの学

■表. IEEE P70xx標準化プロジェクト（敬称略）

	タイトル	役職者	開始/完成目標
P7000	倫理的設計のモデルプロセス	議長：John Havens (IEEEコンサルタント、社会運動家) 副議長：Sarah Spiekermann (ウィーン経済・経営大学)	2016年9月/ 2018年末
P7001	自律システムの透明性	議長：Alan Winfield (西イングランド大) 幹事：江川尚志 (NEC)	2016年12月/ 2018年1月
P7002	データプライバシーのプロセス	議長：Michelle Dennedy (Cisco) 副議長：Aurelie Pols (Mind Your Privacy (コンサルタント会社)) 幹事：Matthew Silveira (Objective Business Solutions)	2016年12月/ 2018年1月
P7003	アルゴリズム的バイアス (差別)	議長：Ansgar Koene (ノッティンガム大) 副議長：Paula Boddington (オックスフォード大)	2017年2月/ 2018年7月
P7004	子供と学生データのガバナンス	議長：Marsali Hancock (DQ Institute、シンクタンク、NPO)	2017年3月/ 2019年2月
P7005	従業員データのガバナンス	議長：Ulf Bengtsson (Sveriges Ingenjörer、スウェーデン大学卒エンジニア協会) 副議長：Christina Colclough (UNI Global Union、労働組合の国際組織)	2017年3月/ 2017年12月
P7006	パーソナルデータAIエージェント	議長：Katryna Dow (Meeco) 副議長：Gry Hasselbalch (DataEthics (シンクタンク)、コペンハーゲン大教授でもある)	2017年3月/ 2017年12月
P7007	用語	議長：Edson Prestes (リオグランデ・ド・スル連邦大学) 副議長：Sandro Rama Fiorini (パリ第12大学)	2017年3月/ 2018年3月
P7008	人を倫理的につき動かすAI	議長：Laurence Devillers (LIMSI、CNRS附属研究所) 副議長：John Sullins (ソノマ州立大学)	2017年7月/ 2018年12月
P7009	AIのフェールセーフ設計	議長：Danit Gal (北京大学、IEEEアウトリーチ委員会議長) 副議長：Alan Winfield (西イングランド大)	2017年7月/ 2018年12月
P7010	AI時代の幸福の指標	議長：John Havens (IEEEコンサルタント、社会活動家) 副議長：Laura Musikanski (Happiness Alliance、NPO)	2017年7月/ 2018年12月
P7011	ニュース源の信頼性の特定と信頼性評価	議長：Joshua Hymen (ピッツバーグ大)	2018年2月/ 2020年4月
P7012	機械可読な個人情報の合意	議長：David Reed (元MITメディアラボ、伝説的インターネット研究者)	2018年2月/ 2019年7月

習用データとして教育現場で作られるデータについて、個人情報保護等の観点から、ユーザがそうした情報にアクセス、収集、保存、利用、共有、破棄する方法を定める。また本標準への適合性要件を定め、ベンダや教育機関が信頼できる機関と出会う方法を定める。

P7005 (従業員データのガバナンス) は、AIの学習用データとして企業が持つデータについてP7004と同様の方法を定める。

P7006 (パーソナルデータAIエージェント) では、現在の集中型の個人情報データ管理が破たんしつつあることを受けて注目を集めている、分散的に保管された個人情報に対する利用申請に対して可否判断を行うAIエージェントについて、エージェントへの入力、学習、倫理、規則、価値を定め、そのAIエージェントの技術的な構成要素を規定する。

P7007 (オントロジー) は、倫理的なロボットやAIに関わる概念や用語を規定する。

P7008 (倫理的誘導) は、AIやロボットは人の感情を動かし行動させる力 (nudge、注意をひくためにひじでそっと突く、少しずつ押すの意味) があるとの認識の下、nudgeを倫理的に行わせることを目指し、nudgeの典型例を記述し、概念を規定し、その機能、必要となる利点を定義し、ロボットやAIの設計での方法論を規定する。

P7009 (フェイルセーフ設計) は、自律・半自律システムで効果的なフェイルセーフ機構を開発、実装、利用するための基本的な方法論を規定する。その一部として、システムがフェイルセーフする能力を計測、試験、認証し、不十分な場合の改善方法を指南する手続きも定める。

P7010 (幸福の尺度) は、来るべきA/ISが普及した世界をより良い場とするためには大局的な指標作りが必要との認識の下、GDPと並ぶ、人間の幸福を測る主観的・客観的データから計算される尺度を規定する。

P7011 (ニュース源の信頼性評価) は、オンラインのニュース提供者の信頼性を半自動で格付けすることを目指し、事実としての正確性を識別し格付けするプロセスを標準化する。アルゴリズムはOSSで定義する。

P7012 (機械可読な個人情報の合意) は、個人が個人情報を提供する条件を記述する方法、及び機械がそれを読み取り同意する方法を標準化する。これにより個人が能動的に条件を企業等のwebサイトに提示し、情報提供やサービスの供与を受けることを可能にする。

以上13個の標準化を大別すると、AIならではの標準はP7001、03、08、個人情報保護系がP7002、04、05、06、12、ソフトウェア系標準の進化系と言えるのがP7000、09、その他がP7007、10、11となる。最多は個人情報保護系の5件となる。AIによるプライバシーの侵害や、不適切な方法で集められたデータがAIの訓練に使われることへの懸念が世界的にも強いことを示しているであろう。

## 4. ISO/IEC JTC1

ISO/IEC JTC1では、1970年代から人工知能の用語定義が進められてきた。まとまった形で発行されたものとしてはISO/IEC 2382-28:1995 'Artificial intelligence-Basic concepts and expert systems' が最初となる。その後機械学習、ニュー



ラルネットワークについても用語が標準化された。現在、これらは全てISO/IEC 2382:2015に統合されている。

近年のAIの盛り上がりを受けてJTC1やISO、IECのTC内では担当分野におけるAIの応用検討が開始されつつある。それら活動の集約点になることを念頭に、2017年10月のJTC1ウラジオストック総会はAIを扱う新Subcommittee, SC42の設立を決議 [JTC1] した。SC名はArtificial Intelligenceで、本SCはJTC1におけるAI標準化を推進し、集約点となり、またJTC1、ISO、IECでのAI応用へ助言することが求められている。

決議はSC42が扱うトピックとしてFoundational standards (基盤的な標準群)、Computational methods (計算法)、Trustworthiness (信頼)、Societal concerns (社会的懸念)を含めることを求めている。またビッグデータを検討してきたJTC1/WG9はSC42へ統合されることとされた。そしてSC42の議長はWael Diab氏 (Huawei、米)、幹事国は米国となった。

SC42の設立と並行してAI関連標準の作成が議論され、現在、2件の標準作成が合意されている。

- ・ Artificial Intelligence Concepts and Terminology (AIの概念と用語、AWI 22989)
- ・ Framework for Artificial Intelligence Systems Using Machine Learning (ML) (機械学習を用いたAIシステムのフレームワーク、AWI 23053)

これらを受けてSC42の第1回会合が2018年4月18～20日に北京で開催された。会合ではSC42の構成が議論され、下記のように1個のWGを設置して前記2個の標準草案の作成を進めるとともに、3個のSGを設置することが決議された [SC42]。倫理や信頼に関係するのは主にSG2となる。

- ・ WG1: Foundational Standards  
コンビーナ: Mr. Paul Cotton (カナダ)  
2個の既存WIを検討
- ・ SG1: Computational approaches and characteristics of artificial intelligence systems  
コンビーナ: Dr. Tangli Liu (中国)  
機械学習や推論など各種のAI技術の特長、画像認識や自然言語処理など各種の用途特化型AIシステムの計算法やアーキテクチャ、産業における具体的応用等を検討して標準化項目を洗い出し
- ・ SG2: Trustworthiness  
コンビーナ: Dr. David Filip (アイルランド)  
透明性、検証可能性、説明性、制御性などを通じたAIに対する信頼確立のためのアプローチ、AIシステムの落とし穴やリスクと対抗策、頑健性や信頼性や正確性やプライバ

シーを達成するアプローチ、AIシステムのバイアス (偏り)の発生源と対策等を探索し標準化項目を洗い出し

- ・ SG3: Use cases and applications

コンビーナ: 丸山文宏氏 (日本、富士通)

AIの応用分野とその利用方法を特定してユースケースを収集し、標準化項目を洗い出し

SC42は今後、年2回のペースで全体会合を開催する。今回は2018年10月にシリコンバレーを予定している。米国はSG2でのバイアスの議論の重要性を訴えており、今後中心な話題の一つになると見込まれる。

なお、本SCに対応するため2018年4月1日付で情報規格調査会配下にSC42専門委員会が設立され、筆者は幹事を拝命している。

## 5. IEC ahG 79

IEC Standardization Management Board (SMB) は2018年2月会合において、自律システムの倫理や信頼、価値においてIECと標準が果たすべき役割を検討するため、ahG 79 (autonomous systems-ethics) の設置を決定した [IEC]。本ahGはIECが取るべきアクションを2018年10月のSMB会合に報告することが求められている。米国と中国を共同議長とし、まずはJTC1/SC42、IECの様々なTC、IEEE、ISO他の標準化団体の関連活動の情報収集とレビューが行われている。日本からは筆者が参加している。

## 6. 通信系の標準化

通信に特化したAIの標準化検討も開始されている。

ITUは2017年6月にXPRIME、多数の国連機関と共同でAI for Goodサミットを開催してハイレベルの対話を行った。第2回は2018年5月に予定されている。またITU-T SG13はFocus Group on Machine Learning for Future Networks including 5G (FG-ML5G) を立ち上げた。これらの詳細は本特集の別稿を参照されたい。

ETSIでは2017年2月にISG Experiential Networked Intelligence (ENI) を、2017年12月にISG Zero touch network and Service Management (ZSM) を設立した。前者は中国系のオペレータとベンダが設立したISGで、SDNやNFVにより仮想化された網資源をAIを組み込んだオーケストレータで制御して網の運用管理の負担を軽減し、また (網の運営管理上の) 新アプリを開発しやすくすることを目指している。後者は世界各国のオペレータとベンダが設立したISGで、ハイブリッドな網での仮想化されたNEやサービスをクロスレイ

やに、e2eで自動化した管理することを目指している。後者の詳細は本特集の別稿を参照されたい。

## 7. AIを巡る標準化での留意点

AIの標準化で、従来と大きく異なる標準は（例えばIoTという巨大分野との比較において）決して多くない。まずアルゴリズム等は当面は競争領域であり、標準化にはなじまない。多くの応用ではAI間の相互接続では必須ではなく、従って標準化は必須ではないし、相互接続する場合でもその手順は現在のところ従来のICTシステムでの各種プロトコルからパラメータが少し変わる程度となりそうである。究極の汎用天然知能(?)である人間のプロトコル（外交儀礼）も概念レベルで見ればICTシステム間のプロトコルと同様であることを考えれば、近い将来において概念レベルの変更はないであろう。前掲の「米国R&D戦略」では様々な標準が列挙されているが、既存標準の拡張や修正で対応できそうなものが多い。

またAIの標準化では、AIという言葉の理解に基づく混乱が往々にして起きる。これは特に大きな枠組みを定める大原則の議論で起きやすい。

AIは様々に定義されている。例えば平成28年度情報通信白書 [noDef] は研究者間のAI観の違いを端的に示している。ISO/IEC 2382:2015 [IsoIec2382] はAIを「推論や学習など一般に人間の知能と結び付けられる機能を実現する能力をもつ機能単位」と定義するが「一般に人の知能と結び付けられる機能」として万人が賛同する定義はない。このため万人が納得するAIの定義は存在せず、その結果、ある特定の技術がAIに含まれるかは自明ではない。大昔に遡れば全てのICT技術は人工知能を目指した営みであったとの声すら聞く。

AI標準化の議論はこうした曖昧さを抱えているが、一般には現在ないし近い将来に実用化されるAI技術、特に現在のAIブームを引き起こしたディープラーニングと機械学習を事実上想定していることが多い。例えばITU-Tで設立されたFG-ML5Gでは、名称にAIではなく機械学習を使っている。これはこうした混乱を避ける知恵であろう。それでもディープラーニング以外にAIに分類され得る技術は多く、これらを全て包含しようとすると記述は往々にして曖昧になりがちである。

これに対し、カーツワイルのシンギュラリティ論 [カーツワイル] ほかに影響された、汎用人工知能 (Artificial General Intelligence: AGI) を想定する場合もある。AGIは2030年前後の実現を目指して世界的に開発競争が行われており、

ディープラーニングをも大きく超えるインパクトを与える可能性を持っている。実現の目途が立ったとは言えない技術であるが、各種基本法の大改正のスパンは一般に数十年であるし子供の平均余命は半世紀以上あるため、法律や法哲学、労働問題や教育問題を考える人々にとって本当に2030年、仮に伸びて2050年であってもAGIが作られるなら、今から議論する価値はある。IEEE EADは汎用人工知能のため1章を割いている。

AI標準を見る際にはこのように時間軸、技術的な性質や性能が大きく異なる議論が混じることに留意されたい。

## 8. おわりに

AIを巡り行われている課題の分析とそれに応じた各国での大きな枠組みの設定、その具体化の一つであるAIの標準化を、特に倫理や信頼 (trustworthiness)、ELSIに注目して紹介した。現在の標準化の多くはAI全てを広く対象としているが、今後は個別分野に特化した標準が、現在の標準を基盤として開発されることとなろう。それら実務に強く影響する標準について今後注意し関与することが重要である。

### 参考文献

- [noDef] 総務省、平成28年度情報通信白書 図表4-2-1-4、2016年7月。
- [カーツワイル] レイ・カーツワイル、ポスト・ヒューマン誕生 コンピュータが人類の知性を超えるとき、NHK出版、2007年1月。
- [IsoIec2382] ISO/IEC, Information technology-Vocabulary, 2015。
- [懇談会] 人工知能と人間社会に関する懇談会、「人工知能と人間社会に関する懇談会」報告書、2017年3月24日。
- [総務省ガイドライン] AIネットワーク社会推進会議、「国際的な議論のためのAI開発ガイドライン案」、2017年7月28日。
- [EU9原則] European Group on Ethics in Science and New Technologies, “Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems”, March 2018。
- [アシロマ原則] Future of Life Institute, “Asilomar AI principles”, January 2017。
- [米国R&D戦略] National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, “The National Artificial Intelligence Research and Development Strategic Plan”, October 2016。
- [JTC1] ISO/IEC JTC 1, Resolutions Adopted at the 32nd Meeting of ISO/IEC JTC 1, 2-6 October 2017 in Vladivostok, Russia, N13630, 2017-10-08。
- [SC42] ISO/IEC JTC 1/SC 42, Resolutions from the Inaugural JTC 1/SC 42 Plenary-Beijing 2018, N 78, 2018-04-21。
- [IEC] IEC SMB, autonomous systems-ethics, decision 161/4, 2018/02。