

第3回 人間中心のAI社会原則検討会

AI技術の方向性及び国際議論

7月5日

日本電気株式会社

永沼美保

Orchestrating a brighter world

未来に向かい、人が生きる、豊かに生きるために欠かせないもの。
それは「安全」「安心」「効率」「公平」という価値が実現された社会です。

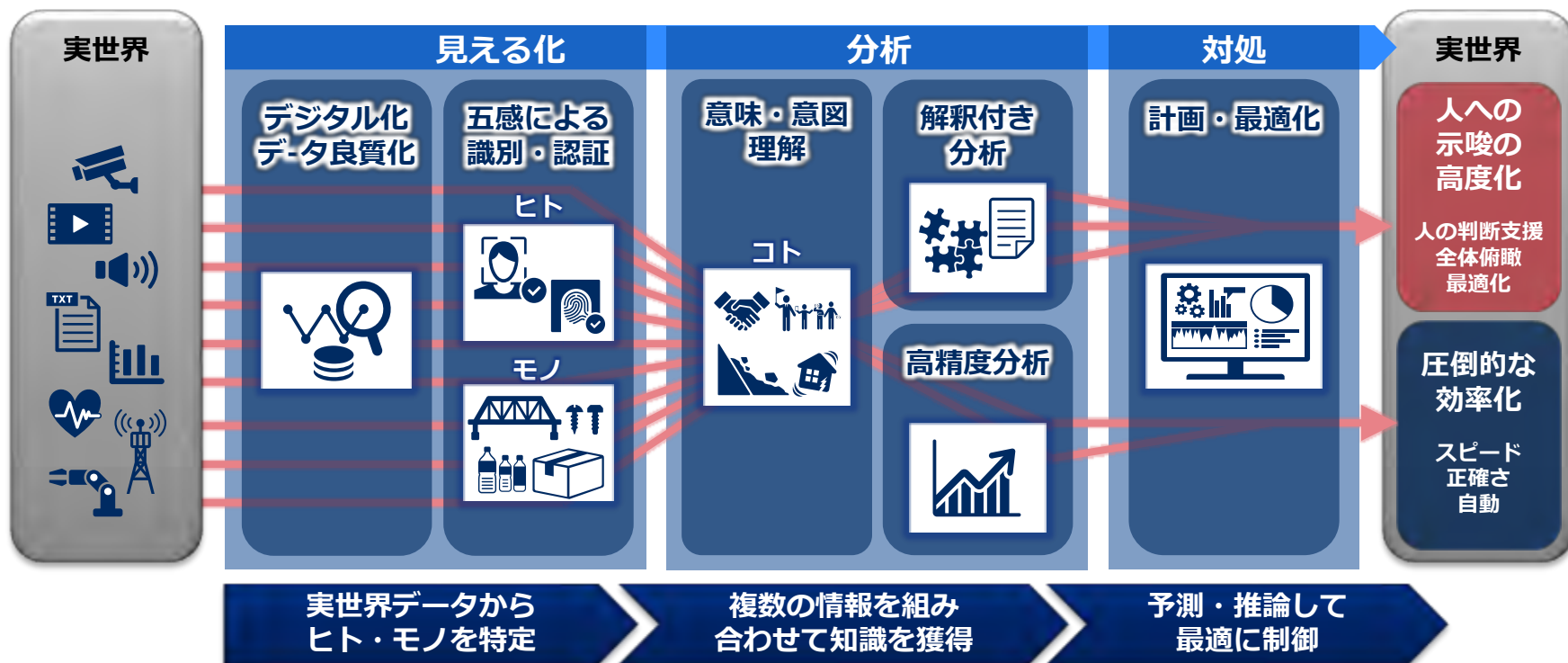
NECは、ネットワーク技術とコンピューティング技術をあわせ持つ
類のないインテグレーターとしてリーダーシップを発揮し、
卓越した技術とさまざまな知見やアイデアを融合することで、
世界の国々や地域の人々と協奏しながら、
明るく希望に満ちた暮らしと社会を実現し、未来につなげていきます。

- 社会課題解決の2つの面に関する議論の切り分け
 - 効率化のための技術革新
 - 知性の高度化、社会的側面を含む議論
- リスクの理解とリスクの低い分野における利用の促進
- 国際議論の活用



実世界のデータを価値へ変換

複数の技術群とデータを組み合わせることで、社会価値の創造を目指す



国際的標準化機関等は機械学習を中心に倫理・リスクを議論

～2016

IEEE EAD
2016/04からAIの問題を分析、
13委員会中12委員会は**機械学習を**
念頭にリスクを洗い出し

重複参加者多数

英国標準協会BS8611
製品安全の概念を応用した
AI・ロボットのリスク分析を
強制力のないガイドラインに

2017

IEEE P7000～P7013
透明性、プライバシー、差別、
フェールセーフなど**機械学習の実務**
上の問題を対象に標準策定開始

情報交換

ISO/IEC JTC 1/SC 42
AIの技術、倫理解決を助ける
技術標準策定を開始

IEC SMB
AI倫理についてIECとしての役割に
ついて検討を開始

ITU AI for good summit
国際的なAI利用の対話の
場を提供

ITU-T FG-ML5G
5Gへの機械学習応用の
技術面を集中討議

2018～

課題を随時増やし
つつ標準策定中

IEEE P70xx標準化プロジェクト一覧 (1/2)

	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	アルゴリズムック 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 Ingenjorer, 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

AIならではの: P7001, 03, 08, 13; 個人情報保護系: P7002, 04, 05, 06, 12; ソフトウェア系: P7000, 09
その他: P7007, 10, 11

開始: 上位委員会 (NESCOM)でプロジェクトが設立承認された日時。完成目標: 親ソサエティでの第1回スポンサー投票目標日時。
投票成功=技術的に完成とみなされる。その後議論の公平性等のチェックを経て半年後に正式標準に

IEEE P70xx標準化プロジェクト一覧 (2/2)

	タイトル	役職者	開始/完成目標
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付属研究所) 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 Hyman (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への合意	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
P7013	自動化された顔分析の包摂とアプリケーション	議長: Joy Buolamwini (MIT助教)	2018年5月/ 2019年7月

AIならではの: P7001, 03, 08, 13; 個人情報保護系: P7002, 04, 05, 06, 12; ソフトウェア系: P7000, 09
その他: P7007, 10, 11

開始: 上位委員会 (NESCOM)でプロジェクトが設立承認された日時。完成目標: 親ソサエティでの第1回スポンサー投票目標日時。
投票成功=技術的に完成とみなされる。その後議論の公平性等のチェックを経て半年後に正式標準に

議論に向けたポイント

- 社会課題解決の2つの面に関する議論の切り分け
 - ・ 効率化のための技術革新
 - ・ 知性の高度化、社会的側面を含む議論
- リスクの理解とリスクの低い分野における利用の促進
- 国際議論の活用



補足資料

P7000:倫理的設計のモデルプロセス

- 委員会名称: Model Process for Addressing Ethical Concerns During System Design (システム設計時の倫理上の懸念を扱うためのモデルプロセス)
- スポンサー(*): Computer Society, Software & Systems Engineering Standards Committee (C/S2ESC)
- 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)
- 開始: 2016/09, 第1回スポンサー投票(**)目標: 2018/07, 投票: 個人
実際には2018年末の完成を目指すことに
- 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出来るよう手順のモデルを定める
- 関連する標準:存在せず。関連文書: IEEE, 及びIEEE Computer Society/ACMソフトウェアエンジニアの倫理綱領 (code of ethics)
- 利用者: 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

* スポンサー: 技術的内容に責任を持つ組織

** スポンサー投票: 承認されると、技術的には完成と見なされる。その後、役員会で手続きの公平性他がチェックされ、問題がなければ約半年後に正式標準に

P7001:自律システムの透明性

■ 委員会名称: Transparency of Autonomous Systems (自律システムの透明性)

■ スポンサー: Vehicular Technology SocietyのIntelligent Transportation Systems (VT/ITS) とRobotics and Automation SocietyのStanding Committee for Standards (RAS/SC) が共同

■ 議長: Alan Winfield (西イングランド大)、幹事: 江川尚志 (NEC)

■ 開始: 2016/12, 第1回スポンサー投票目標: 2018/01, 投票: 個人

■ 透明性のレベルを測定し、試験可能に

- 設計者が標準への適合レベルを評価可能に。「本製品の透明性はレベル3」等
- 透明性向上のための示唆提供。例: センサや内部状態は安全なストレージに保存

■ 背景: 社会にAI/ASが受容されるために透明性が必要に

- ロボットの動作の理由を理解可能とし、予期せぬ動作は問いただし可能にする必要
- 事故発生時の原因究明や証拠提出の必要性

■ 関連する標準: 存在せず

P7002:データプライバシーのプロセス

委員会名称: Data Privacy Process (データプライバシーのプロセス)

スポンサー: Computer Society, Software & Systems Engineering Standards Committee (C/S2ESC)

議長: Michelle Denny (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)

開始: 2016/12, 第1回スポンサー投票目標: 2018/01, 投票: 個人

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

- 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が可能に

対象ライフサイクル: 開発から品質保証、価値実現まで

関連する標準: ISO/IEC JTC 1/SC40/WG1の38500シリーズ、特に38505-1, The application of ISO/IEC 38500 to the governance of data, ISO/IEC 29100, Privacy framework

利用者: 個人情報を含むシステムの関係者

P7003:アルゴリズムミックバイアス

- 委員会名称: Algorithmic Bias (アルゴリズムミックバイアス)
- スポンサー: Computer Society, Software & Systems Engineering Standards Committee (C/S2ESC)
- 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)
- 開始: 2017/02, 第1回スポンサー投票目標: 2018/07, 投票: 個人
- 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
- 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、他を記述することを想定
- 将来的に本標準に基づく認証を想定「本サービスは人種や性別に対する差別がないことをP7003に基づき認証します」
- 関連する標準:存在せず

P7004:子供と学生のデータのガバナンス

委員会名称: Child and Student Data Governance

スポンサー: Computer Society/Learning Technology (C/LT)

議長: 議長: Marsali Hancock (DQ Institute, シンクタンク, NPO)

承認: 2017/03, 第1回スポンサー投票目標: 2019/02, 投票: 個人

● 第1回電話会合: 2017/05/31

AIの学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、ベンダーや教育機関が信頼できる機関と出会う方法を定める。

P7005の子供/学生版。

必ずしもAI向けデータ収集に限られていない。おそらく個人情報保護系の関係者が、この機を捉えて立ち上げた標準化プロジェクト。会合では「カリフォルニア州の法律では」といった発言が飛ぶ。

関連する標準:存在せず

P7005:透明性ある従業員データのガバナンス

委員会名称: Transparent Employer Data Governance

スポンサー: Computer Society/Software & Systems Engineering Standards Committee (C/S2ESC)

議長: Ulf Bengtsson (Sveriges Ingenjorer, 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)

承認: 2017/03, 第1回スポンサー投票目標: 2017/12, 投票: 個人

● 第1回電話会合: 2017/06/12; 実際の完成は2018年の後半か

AIの学習用データとして企業が持つデータについて、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、ベンダーや従業員が信頼できる機関と出会う方法を定める。

P7004の従業員版。GDPRの一具体例と自らを位置づけている

必ずしもAI向けデータ収集に限られていない。おそらく個人情報保護系の関係者が、この機を捉えて立ち上げた標準化プロジェクト。

関連する標準:存在せず

P7006: パーソナルデータAIエージェント

委員会名称: Personal Data Artificial Intelligence (AI) Agent

スポンサー: Computer Society/Software & Systems Engineering Standards Committee (C/S2ESC)

議長: Katriyna Dow (Meeco)

副議長: Gry Hasselbalch (DataEthics (シンクタンク). 兼コペンハーゲン大教授))

承認: 2017/03, 第1回スポンサー投票目標: 2017/12, 投票: 個人

● 第1回電話会合: 2017/06/14

各個人ごとに異なる個人情報保護方針に従って個人情報利用の可否を定めることは、中央集中型のデータ管理では困難となりつつある。このため、データは各所に保存するが、ポリシーはどこか（クラウド上、個人端末等）に保存し、そこが「あなたの〇〇データを使いたい、許可願う」という問い合わせに応答し可否判断する、との手法が注目を集めつつある。

本標準はそうした判断を行うAIエージェントの標準として、個人が制御すべきそのAI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。

関連する標準: 存在せず

P7007:倫理的ロボットと自律システムの用語

- 委員会名称: Ontological Standard for Ethically Driven Robotics and Automation Systems
- スポンサー: Robotics and Automation Society/Standing Committee for Standards (RAS/SC)
- 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)
- 承認: 2017/05, 第1回スポンサー投票目標: 2018/03, 投票: 個人
 - 第1回電話会合: 未定
- 倫理的なロボットやAIに関わる概念や用語を規定
(例えばBS8611:2016ではethical harm, ethical disaster他の用語を定義)
- 関連する標準:存在せず
(実際にはIEEE自身を含め、そうした標準は分野を限れば既に存在する。また他機関とも調整が必要になると見込まれる)

P7008: ロボットやAIによるユーザの倫理的誘導 (nudge)

委員会名称: Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems

スポンサー: Robotics and Automation Society/Standing Committee for Standards (RAS/SC)

議長: Laurence Devillers (LIMSI, CNRS付属研究所)

副議長: John Sullins (ソノマ州立大学)

承認: 2017/07, 第1回スポンサー投票目標: 2018/12, 投票: 個人

● 第1回電話会合: 未定。設立から半年近く経過したが未だに議長団内部で議論中

AIやロボットは人の感情を動かし行動させる力を持つ。これを本標準は' nudge ' (注意をひくためにひじでそっと突く; 少しずつ押す) と呼ぶ。既に chatbot 等では nudge を使った顧客対応や購買誘導の試みあり

AIにこうした力がある以上、それを倫理的に行わせるべき、が本標準の立場。そのために 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

関連する標準: 存在せず

利用者: ロボットの製造業者、サービスやソリューションのプロバイダ、機器のサプライヤとユーザ

P7009: 自律・半自律システムのフェイルセーフ設計

委員会名称: Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems

スポンサー: Reliability Society/IEEE Reliability (RS/SC), Computer Society/Software & Systems Engineering Standards Committee (C/S2ESC)

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)

承認: 2017/07, 第1回スポンサー投票目標: 2018/12, 投票: 個人

●第1回電話会合: 2017/10/16 (準備会合: 2017/09/21)

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。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を含む。

関連する標準:存在せず

利用者: 自律・半自律システム製造にかかわる技術系の企業、エンジニア、開発者、研究者、主管庁および社会

P7010: AI/ASを踏まえた幸福の尺度

委員会名称: Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems

スポンサー: Systems, Man, and Cybernetics Society/Standards Committee (SMC/SC)

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)

承認: 2017/07, 第1回スポンサー投票目標: 2018/12, 投票: 個人

● 第1回電話会合: 2017/09/07

来るべきAI/ASが普及した世界をより良い場とするため、GDPに加え、主観的・客観的データから計算される、人間の幸福を測る尺度を規定

GDPを超える指標、は様々に検討・提案されている。IEEE SAのAI関連活動を実質的に仕切るHavens氏は永年こうした活動に興味があったらしく、氏の想いで設定された標準化プロジェクト。

関連する標準: 存在せず

利用者: 製造業者、サービスやソリューションのプロバイダ、プログラマ、エンジニア、技術関係者

P7011: ニュース源の信頼性の特定と信頼性評価

委員会名称: Process of Identifying and Rating the Trustworthiness of News Sources

スポンサー: Social Implications of Technology/Social Implications of Technology Standards Committee (SSIT/SC)

議長: Joshua Hyman (ピッツバーグ大)

承認: 2018/02, 第1回スポンサー投票目標: 2020/04, 投票: 個人

● 第1回電話会合: ----

ニュース提供者の信頼性を半自動で格付け。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し、それによりオンラインのニュース提供者を格付け。アルゴリズムはOSSで定義。

現在、ネットニュースの格付けの標準は存在せず、その結果フェイクニュースが氾濫。本標準を通じニュースとメディアへの信頼性を取り戻す

関連する標準: 存在せず

関係者: ニュースメディア、ネット検索、ソーシャルメディア、オンライン広告、ソフトウェアエンジニア、OSSコミュニティ、一般消費者、ウェブ開発者

P7012: 機械可読な個人情報の合意

委員会名称: Machine Readable Personal Privacy Terms Recommendation

スポンサー: Social Implications of Technology/Social Implications of Technology Standards Committee (SSIT/SC)

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)

承認: 2018/02, 第1回スポンサー投票目標: 2019/07, 投票: 個人

● 第1回電話会合: ----

個人が個人情報を提供する方法、および機械がそれを読み取り同意する方法を標準化。これにより個人が自己のプライバシーを保つ条件を企業等の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

プライバシーポリシーを定める標準ではない。ポリシーは同意が不要

関連する標準: 存在せず

関係者: 個人、個人情報を取り扱う企業

P7013:自動化された顔分析技術の包摂とアプリケーション

委員会名称: Inclusion and Application Standards for Automated Facial Analysis Technology

スポンサー: Social Implications of Technology/Social Implications of Technology Standards Committee (SSIT/SC)

議長: Joy Buolamwini (MITメディアラボ助教)

承認: 2018/05, 第1回スポンサー投票目標: 2019/07, 投票: 個人

● 第1回電話会合: ----

AI用顔データの多様性の定義を遺伝子の表現型と人口動態の観点から行い、自動的にされた顔分析のための正確性のレポート方法やデータ多様性のプロトコルや規定を確立し、自動化された顔分析技術を使うべきでない文脈を決めるためのレーティングシステムを描写

レポート方法やプロトコルの規定を通じ顔分析技術の透明性を向上

関連する標準:存在せず

関係者: 第一に顔分析技術にかかわる技術者や研究者、これら技術に関わる調達責任者や監査者や意思決定者、一般大衆

■ タイトル: Artificial Intelligence (2017/10設立)

■ 議長: Wael Diab (Huawei), 幹事国: 米国

■ Scope

- 1. Serve as the focus and proponent for JTC 1's standardization program on Artificial Intelligence (JTC1内のAI標準化の集約点かつ推進役)
- 2. Provide guidance to JTC 1, IEC, and ISO committees developing Artificial Intelligence applications (JTC1, IEC, ISOの各委員会のAI標準化をガイド)

■ ITSCJ配下に審議団体として国内専門委員会設置

■ 議長: 杉村(産総研)、幹事: 江川(NEC), 丸山(富士通)

■ 第1回会合: 2018/04/18-20, 北京; 組織構成を決定

- WG1: Foundational Standards (基礎的標準群)
コンビーナ: Mr. Paul Cotton (Canada)
- SG1: Computational approaches and characteristics of artificial intelligence systems (AIシステムへの計算機学的なアプローチと特性)
コンビーナ: Dr. Tangli Liu (China)
- SG2: Trustworthiness (AIに対する信頼)
コンビーナ: Dr. David Filip (Ireland)
- SG3: Use cases and applications (ユースケースとアプリ)
コンビーナ: 丸山文宏 (日本、富士通)

■ 最初のWI: AIのフレームワーク、用語の2件

IEC SMB アドホック (ahG) 79

2018/02, IEC SMB (ボード) はSMB直下にahG 79 (autonomous systems – ethics ad-hoc) を設立

役割

- 特にAutonomous Systems (自律システム) を対象としてIECと標準化が倫理、トラスト、価値について果たすべき役割を評価
- 2018/10のSMB会合にIECが行うべきことを答申

日本からもエキスパートが参加

 **Orchestrating** a brighter world

NEC