

情報技術標準化フォーラム2018

AIを巡るルール形成と標準化の概観

2018/05/22

日本電気株式会社 江川尚志

新しい技術には新しいルール・標準

電話が回線交換（電話交換機） → パケット交換（ルータ他）に

■ 音声に許容される遅延 → 変更せず

■ 品質基準にパケットロス、遅延ジッタ等を追加

■ 新しいアプリを開発できる → APIを新規に規定

■ 等々々々

産官学を集めたフォーラム他で議論

電気通信事業法他の業法や省令、標準をレビューし新規策定、改廃

全ては現場から生まれる
現場を知る人間が政治家や官庁に政策提言し、
新しい時代のルール作りへ貢献するのは
「市民としての義務」

（断じて後ろ暗いロビーイングではない）

AIのルール：超長期視点

(前略)

■ 有機体のロボットが出現し、人間と融合し始めている。これまで人は権利主体でロボットは客体、生物と非生物には境界があるとアприオリに仮定していたが、これはセム系一神教由来の「人は羊飼い」との妄想に過ぎないのでは、を本気で疑うべき時が来た

■ ロボットは制御できるし、するのだ、を前提に現在の法体系は出来ている。だが突然変異し自己増殖する有機体のロボットが本当に制御できるのか。殺虫剤を準備し駄目なら潰す以外の方法はないのではないか。ならば、潰される対象は人間の遺伝子やアルゴリズムを豊富に含んでいる筈なので、法的位置づけを今から考えるべき

(以下略)

2016/11,情報ネットワーク法学会総会にて夏井高人明大教授の基調講演「サイバー法の未来」の一部を筆者が要約

法哲学や憲法・基本法、子供の教育他を考えると
2030年、2050年は遠くない未来

すいません、何仰ってるのか分かりません。
これ、自律システムの話ですよね。
現在のAIって全然自律してないんですが？

ある講演会での、AI研究者からの質問

2018年現在、AIに対する理解の進展とともに
少しずつ現場感覚に近い議論にシフト

AIに対して社会が抱く懸念

内閣府「人工知能と人間社会に関する懇談会
(~2017/3)」の議論のごく一部

Ethics (倫理)

- AIを用いた恋愛ビジネスはどこまで許されるか
- AIに与えるデータを選択することで、最適化とそれによる人間行動の誘導が可能に
- AIによる与信審査の手続きが不透明な場合、人工知能が人間の価値を測ることに

● 人

● L

● 人

● 人

● 人

● 医

- ルールがないことによる萎縮効果

Society (社会) (懇談会では更に経済、社会、教育に分類)

- 労働流動性が高まったときに転職等のコストは誰が支払うべきか
- 従来は多数の部下を使い総合力で進めていた仕事がAIと個人だけで実行可能に
- 人工知能による新サービスや人間の代替が進展すれば、例えばパターンリズムの問題が現実化

自分が関わるルール・標準の寿命や射程を考えつつ
いま必要なものを検討すべき

AIで必要となる従来型・新規のルールと標準

従来型のルールと標準 (2016年10月「米国AI研究開発戦略」より)

- ソフトウェアエンジニアリング：システムの複雑性や持続性、セキュリティの管理、創発的な振る舞いのモニタと制御
- パフォーマンス：正確性、信頼性、頑健性、アクセシビリティとスケーラビリティの保証
- 尺度：パフォーマンスやコンプライアンスに影響する因子の定量化
- 安全性：システム、人間とコンピュータの相互作用、制御系、規制へのコンプライアンスについてのリスク管理とハザード解析を評価
- 使い勝手：インターフェースや制御が効果的で、効率よく、直感的であることの保証
- 相互接続性：交換可能なコンポーネントやデータ、トランザクションモデルを標準や互換インターフェースを通じて規定
- セキュリティ：情報のCIA (秘匿性、完全性、可用性) とサイバーセキュリティを記述
- プライバシー：情報が処理、転送、保存される際の保護を制御
- 追跡可能性：イベント (の実行、試験、完了) の記録、データのキュレーション
- ドメイン：ドメイン特有の標準の用語や対応するフレームワークの定義

AIならではの新規のルールと標準

- 「AIって凄すぎ。不気味、怖い、私は不要になるのでは」
- **人々に使って貰うため、人々の不安を払拭する必要**
 - 倫理 (ethics), 信頼 (trustworthiness), ELSI (Ethical, Legal, Social Issue) をキーワードとするルールや標準が必要に
 - 差別を防ぐ標準、透明性を規定する標準、AI向け個人情報保護 他

各種レベルでのルール検討

大方針 社会合意レベル

■ 総務省 AI開発ガイドライン

■ 欧州委員会 AI, ロボット、自律システムに関する声明

■ アシロマの原則

■ Partnership on AI, Tenet(信条) 等々

AI全般に関する ルール・指針・標準

■ 内閣府人工知能と人間社会に関する懇談会

■ 経団連 AI活用原則TF

■ ISO/IEC JTC 1/SC 42 用語、フレームワーク

■ IEEE EAD 等々

各分野や問題に特化した ルール・指針・標準

■ 知財戦略本部 知的財産推進計画2017

■ 国交省 自動運転における損害賠償責任に関する研究会

■ ETSI ISG Zero touch Service Mngt

■ IEEE P70xx 等々

総務省 国際的な議論のためのA I 開発ガイドライン案

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

1. **連携の原則**; 開発者は、A I システムの相互接続性と相互運用性に留意する

(主にA I システムのリスクの抑制に関する原則)

2. **透明性の原則**; 開発者は、A I 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
3. **制御可能性の原則**; 開発者は、A I システムの制御可能性に留意する
4. **安全の原則**; 開発者は、A I 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
5. **セキュリティの原則**; 開発者は、A I システムのセキュリティに留意する
6. **プライバシーの原則**; 開発者は、A I 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
7. **倫理の原則**; 開発者は、A I 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)

8. **利用者支援の原則**; 開発者は、A I 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
9. **アカウントビリティの原則**; 開発者は、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

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」(2017/07)

欧州委員会 AI, ロボット、自律システムに関する声明

- 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タビリティ)
- 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)

European Group on Ethics in Science and New Technologies, “Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems”, March 2018.

アシロマAI原則

研究課題

1. **研究目標**：研究の目標となる人工知能は、無秩序な知能ではなく、有益な知能とすべきである。
2. **研究資金**：コンピュータサイエンスだけでなく、経済、法律、倫理、および社会学における困難な問題を孕む有益な人工知能研究にも投資すべきである。そこにおける課題として、以下のようなものがある。
将来の人工知能システムに高度なロバスト性をもたせることで、不具合を起こしたりハッキングされたりせずに、私たちの望むことを行えるようにする方法。
人的資源および人々の目的を維持しながら、様々な自動化によって私たちがより繁栄させるための方法。
人工知能に関わるリスクを公平に管理する法制度を、その技術進展に遅れることなく効果的に更新する方法。
人工知能自身が持つべき価値観や、人工知能が占めるべき法のおよび倫理的な地位についての研究。
3. **科学と政策の連携**：人工知能研究者と政策立案者の間では、建設的かつ健全な交流がなされるべきである。
4. **研究文化**：人工知能の研究者と開発者の間では、協力、信頼、透明性の文化を育むべきである。
5. **競争の回避**：安全基準が軽視されないように、人工知能システムを開発するチーム同士は積極的に協力するべきである。

倫理と価値

6. **安全性**：人工知能システムは、運用寿命を通じて安全かつロバストであるべきで、適用可能かつ現実的な範囲で検証されるべきである。
7. **障害の透明性**：人工知能システムが何らかの被害を生じさせた場合に、その理由を確認できるべきである。
8. **司法の透明性**：司法の場においては、意思決定における自律システムのいかなる関与についても、権限を持つ人間によって監査を可能とする十分な説明を提供すべきである。
9. **責任**：高度な人工知能システムの設計者および構築者は、その利用、悪用、結果がもたらす道徳的影響に責任を負いかつ、そうした影響の形成に関わるステークホルダーである。
10. **価値観の調和**：高度な自律的人工知能システムは、その目的と振る舞いが確実に人間の価値観と調和するよう設計されるべきである。
11. **人間の価値観**：人工知能システムは、人間の尊厳、権利、自由、そして文化的多様性に適合するように設計され、運用されるべきである。

12. **個人のプライバシー**：人々は、人工知能システムが個人のデータ分析し利用して生み出したデータに対し、自らアクセスし、管理し、制御する権利を持つべきである。
13. **自由とプライバシー**：個人のデータに対する人工知能の適用を通じて、個人が本来持つまたは持つはずの自由を不合理に侵害してはならない。
14. **利益の共有**：人工知能技術は、できる限り多くの人々に利益をもたらし、また力を与えるべきである。
15. **繁栄の共有**：人工知能によって作り出される経済的繁栄は、広く共有され、人類すべての利益となるべきである。
16. **人間による制御**：人間が実現しようとする目的の達成を人工知能システムに任せようとする場合、その方法と、それ以前に判断を委ねるか否かについての判断を人間が行うべきである。
17. **非破壊**：高度な人工知能システムがもたらす制御の力は、既存の健全な社会の基盤となっている社会的および市民のプロセスを尊重した形で改善に資するべきであり、既存のプロセスを覆すものであってはならない。
18. **人工知能軍拡競争**：自律型致死兵器の軍拡競争は避けるべきである。

長期的な課題

19. **能力に対する警戒**：コンセンサスが存在しない以上、将来の人工知能が持ちうる能力の上限について強い仮定をおくことは避けるべきである。
20. **重要性**：高度な人工知能は、地球上の生命の歴史に重大な変化をもたらす可能性があるため、相応の配慮や資源によって計画され、管理されるべきである。
21. **リスク**：人工知能システムによって人類を壊滅もしくは絶滅させうるリスクに対しては、夫々の影響の程度に応じたリスク緩和の努力を計画的に行う必要がある。
22. **再帰的に自己改善する人工知能**：再帰的に自己改善もしくは自己複製を行える人工知能システムは、進歩や増殖が急進しうるため、安全管理を厳格化すべきである。
23. **公益**：広く共有される倫理的理想のため、および、特定の組織ではなく全人類の利益のために超知能は開発されるべきである。

日本政府の各種委員会報告

データと競争政策に関する検討会	2017年6月	公正取引委員会
人工知能と人間社会に関する懇談会	2016年3月	内閣府
国際的な議論のためのAI開発ガイドライン案	2017年7月	総務省
新たな情報財検討委員会	2017年3月	内閣府
知的財産推進計画2017	2017年5月	知的財産戦略本部
産業構造審議会 知的財産分科会 営業秘密の保護・活用に関する小委員会 第四次産業革命を視野に入れた不正競争防止法に関する検討（中間とりまとめ）	2017年5月	経済産業省
自動運転における損害賠償責任に関する研究会	2017年9月	国土交通省
自動運転の段階的実現に向けた調査研究	2017年3月	警察庁

■ 日本政府のAI検討は欧米に比べても全般的に先行
 ■ 一部委員会は継続され、フェーズ2として議論中

個別のAI標準化活動 (JTC1, IEC, IEEE)

ISO/IEC JTC 1/SC 42 (AI)

■ 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)

■ 第1回会合: 2018/04/18-20, 北京; 組織構成を決定

- WG1: Foundational Standards
コンビーナ: Mr. Paul Cotton (Canada)
- SG1: Computational approaches and characteristics of artificial intelligence systems
コンビーナ: Dr. Tangli Liu (China)
- SG2: Trustworthiness
コンビーナ: Dr. David Filip (Ireland)
- SG3: Use cases and applications
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

共同議長: 米国、中国

- JPからは江川 (NEC) が参加へ

役割

- 特にAutonomous Systems (自律システム) を対象としてIECと標準が倫理、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
行うべき、が認められれば同テーマについてStudy Group (1年間) を設立し議論、その後Technical Committee (常設機関) を設立

IEEEにおけるAI倫理の標準化

2016年4月テーマ洗い出しを開始。合意したものから具体的標準化開始

論点・標準化テーマを洗い出し

Global Initiative on Ethics of Autonomous and Intelligent Systems

一般原則

AI, ASへの価値組み込み

倫理的研究・設計を導く方法論

汎用AIと超知性の安全性と受益者

.....

.....

計18委員会（技術討議は13委員会）

特定された論点とそれに対する勧告はEADという文書に



具体的な標準の策定

P7000:倫理的設計のモデルプロセス

P7001:自律システムの透明性

P7002:データプライバシーのプロセス

P7003:アルゴリズムバイアス

.....

.....

P7012まで計13プロジェクト

EAD: IEEE SAによるAIを巡る論点整理文書

■ 名称 : Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems (EAD, 倫理的整合のとれた設計: AIとASにおいて人の幸福を優先するためのビジョン)

■ 内容 : イニシアチブ各委員会 (の分野) 毎に論点を数点挙げ、その論点の背景、勧告案、参考文献を記述

- (例): 自律システムへの価値組み込み委員会

論点例 : 組み込むべき価値は往々にして特定のユーザーコミュニティや課題に依存

勧告案 : 万人向けの規範や価値のセット構築は非現実的と認め、特定のコミュニティや課題向けの規範セットを特定すべき

- 標準化以外 (教育や立法他) で解決すべき論点も含めてバランス良く広範に論点整理

■ 第1版 (for public discussion版, EADv1)

2016年12月発行; 全136ページ

■ 第2版: 2017年12月発行; 250名の参加によりEADv1を改版

- 基本的には電話会合で各委員会が担当分を改版

- 2017/06のSEAS (Symposium on Ethics of Autonomous Systems) 会合で詰める

- 主な相違点 : 第1版に間に合わなかった and/or 存在しなかった5委員会の内容追加、各委員会の内容精査、拡充

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; 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

AIならではの: P7001, 03, 08; 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

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

今後の予定 (2017/09議長訪問時の情報): 飛行機のフライトレコーダーに倣い、AIの入出力やステートを標準化して記録し、事故調査委員会が調査可能とすることを法的に担保すべき。これをP7001.xまたはP7011 (その時点ではP7010まで立上済) として立ち上げたい

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 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)

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

P7003 標準ドラフト現状

最新版: 2018/02/26会合版。全31ページ

意図せざるbiasがシステムに入ることを防ぐ方法論として下記を記述

- 原則に則った方法論（プロセス）、関係者（人）、利用アルゴリズムを決め正当化する目標（目的）、製品で実現されていることの検証（製品）を助けるガイドラインを提供
- Bias評価のため立ち止まるべき時を実務的に指摘
- Bias品質評価のための検証用データセット選択の基準と手順
- アルゴリズムの限界を設定し周知する方法
- ユーザーの期待を制御する方法
- 用語

これら記述により、システム設計を助け、外部監査を可能とする。なお運用上意図的にbiasが入ることは許容

現在はコンセプトの記述を固めている段階。具体的な方法論や注意すべき事柄の記述は、担当者は指名済だが議論は進んでいない

現在7個のユースケースをベースに議論

人種差別をするchatbot、医師の判断前の子供の感情読み取り、コメント欄の有害性判定、美人コンテスト、アルゴリズムによるコンテンツの改変、パフォーマンス判定アルゴリズムによる教師解雇、犯罪常習者判定の人種差別

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: 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))

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

おわりに

AIのルールや標準の議論が進行中

- 2018年現在、AIの現状を踏まえた、地に足の着いた議論へとシフト中
- 社会合意的なハイレベルの議論から、個別分野、個別の課題ごとの議論まで

AI標準の特徴：倫理、信頼 (trustworthiness), ELSI (ethical, legal, societal issues) 等を扱う標準群の存在

- AIに対する人々の不安を払拭し、使って貰うための標準
- 品質や相互接続など従来の標準も並行して議論

JTC 1/SC 42, IEC, IEEEなどで議論が進行中

- ELSI系はIEEEが先行