

人間中心の AI 社会原則会議（令和 2 年度 第 2 回）議事要旨

1. 日 時 令和 3 年 3 月 24 日(水) 15:00-17:00

2. 場 所 中央合同庁舎 8 号館 5 階 共用会議室 C

3. 出席者※敬称略

議長 須藤 修	中央大学国際情報学部 教授 東京大学大学院 特任教授
安宅 和人	慶應義塾大学 環境情報学部 教授、ヤフー株式会社 CSO
岩本 敏男	株式会社エヌ・ティ・ティ・データ 相談役
浦川 伸一	損害保険ジャパン株式会社 取締役 専務執行役員
江間 有沙	東京大学未来ビジョン研究センター 特任講師
大屋 雄裕	慶應義塾大学法学部 教授
金井 良太	株式会社アラヤ 代表取締役 CEO
國吉 康夫	東京大学大学院情報理工学系研究科 教授 次世代知能科学研究センター センター長
近藤 則子	老テク研究会 事務局長
武田 晴夫	株式会社日立製作所 技師長
中川 裕志	理化学研究所革新知能統合研究センター チームリーダー
永沼 美保	日本電気株式会社デジタルトラスト推進本部 シニアマネージャー
新居日南恵	株式会社 manma 代表取締役社長
羽鳥 裕	公益社団法人日本医師会 常任理事
原田 達也	東京大学先端科学技術研究センター 教授
樋口 知之	中央大学 AI・データサイエンスセンター 所長 理工学部経営システム工学科 教授
平野 晋	中央大学国際情報学部 学部長・教授
福田 剛志	日本アイ・ビー・エム株式会社 理事 東京基礎研究所 所長
松尾 豊	東京大学大学院工学系研究科 教授 日本ディープラーニング協会 理事長
山川 宏	全脳アーキテクチャ・イニシアティブ 代表
吉瀬 章子	筑波大学システム情報系 教授

4. 議題

- (1) 最新の AI に関する動向（研究開発・ELSI）について（有識者プレゼン）
- (2) 意見交換
- (3) その他

5. 資料

- 資料 1 AI 研究の現状について
資料 2 AI エージェントの法的位置づけ
参考資料 1 人間中心の AI 社会原則会議（令和 2 年度 第 1 回）議事要旨

6. 議事要旨

- (1) 最新の AI に関する動向（研究開発・ELSI）について

国内外における最新の AI に関する動向として、原田構成員から資料 1 に基づき AI の研究開発について、中川構成員から資料 2 に基づき AI の ELSI に関するトピックスとして AI エージェントの法的位置づけについて説明が行われた。

- (2) 意見交換

各発表を踏まえ、人間中心の AI 社会原則会議で議論すべき論点、議論にあたって意識しておくべき留意点等について、全体での意見交換が行われた。

(以下はその中での意見の例)

【研究開発】

- AI に関する出版物や論文のインパクトが大きくなる傾向にある中、機械学習に関するトップカンファレンス NeurIPS での日本機関のプレゼンスは低く、アメリカと中国の企業が席卷。国際会議で取り扱われるトピックには変化が見られ、2020 年の CVPR では、Explainable AI、Fairness, Accountability, Transparency and Ethics (FATE)の 2 つが加わった。
- 深層ニューラルネットワーク関連の研究で最もインパクトを与えたものはトランスフォーマー。この出現により自然言語処理の性能が飛躍的に向上した。これをベースとした技術の主なものに OpenAI の GPT と Google の BERT があり、前者の現行バージョン GPT-3 は文章の自動生成で世間に驚きを与えた。またこれらの技術は、自然言語処理のみならず画像分野にも応用が可能で、特に Google の Vision Transformer は畳み込みニューラルネットワークに取って代わる可能性を持っているとされる。
- 自動機械学習 AutoML-Z も深層ニューラルネットワーク関連の研究におけるビッグリープ。行列とベクトルの掛け算、ノルムの計算等の基本的な数学的なオペレーションだけを与え、学習のデファクトとして使われている Stochastic Gradient Descent と同じ計算方法が自動的に導かれた。
- 上記 2 つの研究は計算コストが膨大であり、通常規模の研究機関では対応が難しいが、国内では双曲空間での機械学習手法の研究、説明可能な AI の研究などの分野で研究が進められている。
- AI 研究をけん引してきた IMAGENET がその役割を終えつつある。1 枚の画像に複数のラベルを付ける動きがあること、圧倒的に多量な画像から構成されるデータセットを用いた研究により、IMAGENET が小規模データと呼ばれる状況となったことの 2 点はその理由。また、ラベルのな

いデータを用いて学習する Self-Supervised Learning も話題となっており、この適用がトレンドになりつつある。

- Folding@home の例のように、個々の PC の余っている CPU や GPU のリソースを使うことで高速な演算を行うことが可能。スーパーコンピュータよりも大きなリソースが利用可能な状況にあると言え、今後エッジコンピューティングとこれらの連携で大規模計算を可能にしていくことが重要になってくる。
- プライバシーの問題はゲームチェンジャーとなり得る可能性を秘める。高精度なモデルがあっても、あるタイミングで、学習に用いたデータセットの内容に問題があると判断され、これを取り除くことができなければ、運用していたモデルやサービスが全て駄目になる。
- 個別最適化だけではなく、全体最適化に目を向けることも重要。今後プラットフォーム非依存、分散計算、非同期計算、プライバシー保護、分散データ、多様性、連合学習等も重要になってくるのではないかな。

【ELSI】

- AI エージェントは個人を代理する AI で、これが AI の重要な応用であるということは IEEE の Ethically Aligned Design でも言及されている。AI エージェントの法的位置付けについては、人間と同じ人格を持つ議論は棚上し、ある種の法人格のようなものを持たせるのが現実的ではないか。一方、AI エージェントの決定が人間の社会に実際に影響を与えるようになったとき、どのような問題が起きるかは今後考えていくべきではないか。
- AI エージェントの法的位置づけを肯定的に捉えようとするものに、Lawrence B Solum の論文(1992、2019)、Gunther Teubner の論文(2007)、Burkhard Schafer の論文(2010)、Chopra and White の書籍(2011)、Ugo Pagallo(2013)の書籍等があり、これらの中では、AI に与える法人格について、アクター理論の中に AI エージェントを位置づける取組、AI/AGI の法的格付け別の責任の考え方等が論じられている。
- AI エージェントと人間の間には相互作用があり、AI エージェントが賢くなるにつれ、人間もこれに適応能力が変化する現象が顕著になっている。互いに影響を与え合う状況での AI の法的位置づけの構築が重要ではないか。
- ルールで固定するのではなく、アジャイルにゴールも変化させるアジャイルガバナンスに AI エージェントが適用され、企業活動に影響を及ぼすようになると、AI エージェントの役割はより大きくなる。特に AI エージェントに判断を行わせる場合、AI エージェントは何らかの法的位置付けを持つべきではないだろうか。
- AI エージェントに法人格を持たせるプラクティカルなものには、自動運転車や AI トレーダーが考えられ、自動運転中の事故に対する責任、トレードによるクラッシュの責任をどう落とし込むか、こういった議論が今後必要であろう。（注：AI トレーダーに関しては次の報告書が公開されている アルゴリズム・AI の利用を巡る法律問題研究会. 投資判断におけるアルゴリズム・AI の利用と法的責任. 金融研究 第 38 巻第 2 号, 2019)

【その他】

- プライバシー保護についての考え方、対応の差が AI のパフォーマンスへ影響を及ぼし、ある種の政治的イシューになりかねないとの懸念がある。同時に AI は信用が第一であるとの見方もあり、プライバシー保護のあり方への意識は高まっている。
- AI 研究のブルーオーシャンを探すのは大変な状況だが、双曲空間等のユークリッド空間とは異

なる空間でのニューラルネットワークには全く違う方法論が必要であり、ブルーオーシャン的である。また、よりドメインスペシフィックな分野もこれに類する。

- プライバシーとフェアネスについては、プライバシーを自己決定権の延長、あるいは同意権だと考えると、明確に衝突することが分かっている。個々人の同意に還元しない形でデータ利活用を進めるデータ共同利用権という発想の背景にはこの問題がある。この利用においても自己に影響のない形でデータ利活用されることの保障が必要であり、自己に敵対的に使われないことを保障する法制度か技術的なシステムを構築することが一つの課題ではないか。
- 法人格を AI に認めるかどうかの議論は、利便性や便宜性の問題として考えた方がよい。その際、各国の法制度の違いに意識的である必要がある。
- 賠償責任には民事法的にまさに賠償するという意味と、抑止力という意味がある。このバランスを考え、原因を作った人に萎縮してもらうことが必要かどうか、これが一つの観点になるのではないか。また、民事賠償責任での訴訟には、敗訴するというラベリング効果があり、この効果は非常に大きい。訴訟を使わないで自動的にお金が被害者に出る制度も併せて考える必要があるのではないか。
- AI に対する法人格付与については、株式会社に法人格が与えられた経緯も一緒に考えていくことが有効ではないか。また刑事責任を課す場合の狙い（AI の背後にいる自然人をどうカウントするのか）も検討の対象となるのではないか。
- 一企業の取組という面で一定の限界を感じており、企業間でデータの共有を図り、AI の高度化を進めたいが、個人情報に対するセンシティブティの壁があり企業間のデータの共有が進まない。説明可能な AI 等の技術の進展があっても、データの壁がブレーキになりかねないという懸念がある。この点は今後の議論のポイントになるのではないか。
- 本会議の議論はどういう方向に向くべきかと考えたときに、国外有識者からガラパゴス化しているとの見方をされており、悩ましいと感じている。AI エージェントのような議論は海外では若干の抵抗力的なものもありつつ、ヨーロッパ・アジアともある程度やっていけるというようなプレーヤーとしての立ち回りが国際的に期待されている一方で、FAT 的な対応をどうすれば良いか国内企業にはまだ戸惑いがある。この接続がない中でどうバランスを取っていくのが課題ではないか。
- 日本的な観点や技術を国際的に発信していくやり方があると考えられ、ここから何かの原則を導き出していく方法も考えられる。この場合、国内の企業が置いてきぼりになる懸念もある。一方、世界経済フォーラムや GPAI では国内の企業・プレーヤーをエンカレッジする動きもある。いくつかの道筋を考え、どの方向に議論を進めていくべきか議論すべきではないか。
- AI 活用について、政府、行政が積極的に活用する議論が始まっており、我が国の政府でも AI をどう活用するかという議論が一部既になされている。政府及び全自治体がどう AI と向き合うかを重要なテーマとしてこの会議で議論する旨、トランスペアレンシーの観点からも各省庁に協力願う旨、議長から発言があった。

(3) その他

事務局から今後の予定について説明があった。

以上