

## AI ネットワーク社会推進会議

### AI ガバナンス検討会

#### 第1回 議事概要

#### 1. 日時

平成 30 年 11 月 29 日 (木) 15 : 00～17 : 30

#### 2. 場所

中央合同庁舎 2 号館 10 階 共用 1001 会議室

#### 3. 出席者

##### (1) 構成員

平野座長、江間構成員、大屋構成員、木谷構成員、(代理：株式会社 NTT データ 遠藤 技術開発部 Technology Strategist)、木村構成員、久世構成員 (代理：日本 IBM 株式会社 立花 東京基礎研究所 AI 担当シニアマネージャー)、榊原構成員、三部構成員、城山構成員、杉原構成員 (代理：グーグル合同会社 山口 公共政策担当アナリスト)、鈴木構成員 (代理：株式会社日立製作所 城石 研究開発グループ技術戦略室技術顧問)、高橋構成員、武田構成員、中川構成員、原構成員 (代理：富士通株式会社 中条 デジタルサービス部門 エグゼクティブディレクター)、堀構成員

##### (2) 総務省

鈴木総務審議官、安藤大臣官房総括審議官、井上情報通信政策研究所長、  
香月情報通信政策研究所調査研究部長、向井情報流通行政局情報通信政策課統括補佐  
市川情報通信政策研究所調査研究部主任研究官  
高木情報通信政策研究所調査研究部主任研究官  
小川情報通信政策研究所調査研究部主任研究官

##### (3) オブザーバー

須藤東大教授 (AI ネットワーク社会推進会議議長)、  
実積中央大教授 (OECD デジタル経済政策委員会 (CDEP) 副議長)、  
消費者庁、個人情報保護委員会、文部科学省、経済産業省、情報通信研究機構、  
科学技術振興機構、理化学研究所、産業技術総合研究所

#### 4. 議事概要

##### (1) 運営方針

資料 1 の AI ガバナンス検討会運営方針の確認が行われた。

##### (2) 事務局からの説明

事務局より、資料 2 - 1 に基づき今後の検討事項及び検討の進め方等について、資料 2 -

2に基づき国際的な議論及び海外の議論の動向について、机上資料1に基づき、構成員の皆様からの意見について報告があった。

### (3) 有識者からの発表

資料3から資料5に基づき、以下の発表があった。

- 「AI ネットワーク社会の責任論・利活用の主体論・人間の介在」(松尾 剛行 弁護士)
  
- 『AIの説明』の現状とこれから」(大阪大学 産業科学研究所 原 聡 助教)
  
- 「AIによる意思決定の公平性」(東京大学 生産技術研究所 小宮山 純平 助教)

### (4) 意見交換

<事務局からの説明について主な意見>

【久世構成員(代理:日本IBM 立花 東京基礎研究所 AI担当シニアマネージャー)】

- ・ AIの定義の議論が不十分であると感じる。場合分けをして議論をするべきではないか。

【平野座長】

- ・ AI開発ガイドライン案作成に係る議論において一応AIの定義をして、それをそのままこの利活用ガイドライン案に採用している。AIの定義は100人いれば100通り存在するとも言われる。引き続き皆さんで考えていきたい。

<松尾弁護士からの発表に関する主な意見>

【中川構成員】

- ・ 注意喚起文書等に法的効力があるというお話があったが、例えば海外の公的ではない学会が出したスタンダードはどのくらい効果があるのか。
- ・ 消費者的利用者であっても自分のデータをインプットしてAIの動きを変えることができ、ミスユーズしてしまうことが問題視されている。そこで、問題を回避するためのAIエージェントのようなものが重要になるのではと言う議論がIEEEではされている。

【桃尾・松尾・難波法律事務所 松尾弁護士】

- ・ 日本の裁判所が過失や注意義務違反等を認定する時には、一般的にその立場にいる者はそういう状況が起きたら一体何をすべきなのかということを考えて判断する。外国のものについても「国際的なデファクトスタンダードになっている」ということが言えれば、十分に法的判断に影響すると考える。
- ・ ご指摘の通りだと思うが、できるだけミスユーズをしないように警告してあげるとか、ミスユーズした場合に動作しないような設計をして対応するなどが考えられると思っている。

【堀構成員】

- ・ (SQL インジェクションの対策のような) 対応が必然となる事例と、対応が非常に難しい事例の区別は専門家のみが判別できるものであり、ガイドラインの記述だけでは判別が難しい。今後ガイドラインに加えて、例えば航空機事故の事故調査委員会のような専門家組織の検討をどこかで始める必要があるのではないかと。

#### 【武田構成員】

- ・ 事故調査委員会のような枠組みに加え、AI に関しては、問題が大きくなるために対処する仕組みについての議論もあってよいと思う。文書化しただけでは問題は解決しないので、出口として、そのような活動を国内でやっていくべきだという指針を作るのがよいのではないかと。

#### 【木谷構成員 (代理：NTT データ 遠藤 Technology strategist)】

- ・ オープンソースコミュニティの中で、AI ソフトウェアのバージョンアップや、維持管理終了後などのタイミングで、誰も面倒を見ないケースも存在すると思う。そのようなときに開発者、提供者、使用者の責任という概念だけで整理しきれないケースがあるのではないかと。

#### <原助教からの発表に関する主な意見>

##### 【堀構成員】

- ・ 研究の方向性と世の中のニーズが合致していない可能性が高いのではないかと。  
explainable AI については「説明」で終わっているケースが多いが、その次のステップが必要ではないかと。説明を使って責任を取る、データを修正するなどのサイクルが回ることで重要であり、役に立ったという事例が必要だと思う。なお、技術的には説明可能性を高めると推定精度が落ちるトレードオフの問題もあり、使用目的により何を選択するかも変わってくると思う。

##### 【中川構成員】

- ・ 説明が必要となるケースとして、不安の解消だけでなく、(自分が悪くないことを示すための) エビデンスとして使うケースもあるのではないかと。
- ・ 他方、「みんなが使っている」から「安心」につながるというのは、大変示唆に富む表現だと思う。これは、たくさんのチェックやテストケースを経ているということと等価である。AI のチェック (テストの方法) まで含めてトータルに AI というシステムを考えていくべきではないかと。

##### 【須藤オブザーバー】

- ・ 不安の解消に対しては、いわゆる「AI による説明」も重要だが、「ソーシャルインパクト」も重要ではないかと。一般の利用者としては、「自分の仕事がどうなるのか」や「事故が起こったら誰が責任を取るのか」などの KPI がたくさんある中で、「ここまでは大丈夫」という説明も必要になると思う。

【武田構成員】

- ・ 説明可能性とアカウントビリティの関係についての議論や研究はあるか。

【大阪大学産業科学研究所 原助教】

- ・ それをパーフェクトにやっているものはなく、どちらかというと説明の仕方が中心。ただし、幾らでも間違った説明を出すことができるため、間違った説明でみんなを納得させても、それで説明責任を果たしたと言えるかは微妙だと思う。

<小宮山助教からの発表に関する主な意見>

【堀構成員】

- ・ メカニズムデザインについて、オークションのルール作りなどでは数学的に理論が発達していると言われるが、それとマシンラーニングの分野を繋ぐような研究は生まれているか。

【東京大学生産技術研究所 小宮山助教】

- ・ 国際会議 ICML2018 のベストペーパーが、「こういうポリシーを敷くと次に起こることがどうなるか」というある種のメカニズムデザインの研究であった。
- ・ メカニズムデザインでよく出てくるのはオークションとマッチングだと思うが、オークションはお金が入ってくる判断で、マッチングはお金が入ってこない判断。雇用は契約時にはお金が直接入ってこないのので、企業側と人側のマッチングになる。

【大屋構成員】

- ・ ローンの融資可否などの公平性の問題に対して、ごく一部の難しいケースや微妙なケースで、説明できるようなルーティンを回せるようにするなどの仕組みを考えることが、恐らく社会的には有益なのではないか。

【東京大学生産技術研究所 小宮山助教】

- ・ ローンの話は比較的区別が正当化しやすい。営利企業は絶対返せない人にはローンを貸せない。一方、大学入試は多様性が必要ということもあり必ずしも成績順に取ることが良いわけではない。公平性の厳しさも違うのかもしれない。

【大屋構成員】

- ・ 人間の意思決定であっても、何に対して説明責任を追っているのかということはどうも少し慎重に見るべき問題である。人間が行う入試であっても、結果に対して一切説明しない。そこはアカウントビリティを要求されないということになっているので、AIに何らかの問題があるとしても、例えば非劣勢であれば十分であろうというのは一つの基準として言える。それ以上にどこまで道徳的に考慮するかは各主体のポリシーの問題。

#### 【中川構成員】

- ・ ローンや就職の話になると、人間を雇うのか、AIを雇うのかという判断がどこかで出てくる気がする。それに対する説明をどうするべきかということも考えざるを得ないのではないか。

#### <その他主な意見>

##### 【実績オブザーバー】

- ・ AI自身も自分で説明できないものがある場合、人が介在すれば免責できるのか。例えば囲碁で人間には理解できない手を打ったが、結果的には良かったというものに関して、人間の説明責任をどこまで求めるのかということに疑問がある。人間の関与はどこまで厳密でないといけないとお考えなのか御意見を伺いたい。

##### 【桃尾・松尾・難波法律事務所 松尾弁護士】

- ・ 侮辱や名誉棄損による本人に対する支障が大きくなればなるほど、「AIが判断したのでその理由については説明できない」では済まないという程度も高まってくると思う。
- ・ どこかで線引きをする必要がある。およそ全てのものに人間が介在して説明することが法的に求められるとは思わないので、ベストプラクティスないしガイドラインとして「こういう場合については人間が介在した方が良い」ということを議論することが必要。また、それとリンクするけれども違う問題として、「法律上それをしないと違法」というラインもどこかにあると思う。

##### 【大阪大学産業科学研究所 原助教】

- ・ AIが、人間が予想しない素晴らしいものを出してくれる世界が来るという前提がそもそも成り立つのかが分からない。機械学習を使って何かが出たという時に、当然の結果が出ると「当たり前」と言われて、誰も予想しない結果が出るとみんなの思考が止まってしまう。AIがどうこうというよりも、それを受け入れる人間側の準備が全く整っていない気がする。

##### 【東京大学生産技術研究所 小宮山助教】

- ・ 公平性指標について、どういうものを公平と見なすかということは、法律や世論、つまり人間が決めるものである。
- ・ AIの判断は、豊富なデータのチューニングをすることが上手いが、チューニングをすればするほど説明能力が下がる気がする。例えば、アルファ碁は個々のソースコードは分かるが、要素が100万個あると足し算しても良くわからない。

##### 【堀構成員】

- ・ アルファ碁は自身が打った手について説明ができないわけだが、囲碁のプロ棋士がその手を分析し新しい定石として登録するような話が始まりつつある。
- ・ そういう意味で、AIの後付け説明を行うというタイプの説明の研究もある。後付け説明で

許される問題と許されない問題の両方があると思うが、囲碁や将棋など世界が広がるという場合には後付け説明でも良いのだろう。その辺りの切分けが必要になってくると思う。

以上