

特集 「2016年度人工知能学会全国大会 (第30回)」

## 公開討論「人工知能学会倫理委員会」

江間 有沙 (東京大学), 長倉 克枝 (科学ライター), 武田 英明 (国立情報学研究所)

### 1. 公開討論開催趣旨

2016年6月6日, 昨年度に続き本学会倫理委員会は公開討論を開催した. 19時20分開始という夜遅いセッションであったにもかかわらず, 会員・非会員を含め多くの方々にご参加いただいた.

人工知能がブームとなっている現在, 国内外でさまざまな機関が提言や提案を出している. このような中, 倫理委員会は「人工知能研究者の倫理綱領 (案)」を提示した [人工知能学会倫理委員会 16]. 公開討論の目的は, この綱領案を材料として, どのような条項が必要か, どのような表現が適切か, またそもそも倫理綱領という形式が最良なのかの議論を行うことであった.

公開討論は二部からなり, 第一部では, 人工知能と社会をめぐる状況や倫理委員会の活動を報告した. また, 倫理綱領案の構成や成り立ち, それに対してゲストの土屋 俊氏 (大学評価・学位授与機構) からコメントをいただくなど議論を行うための情報を提供した. 第二部では倫理委員会のメンバと土屋氏で討論を行った. 当日は, 会場の意見を倫理委員会のホームページで募集し適宜取り上げた. また会場からの質問に対しては, 後日ホームページにて登壇者らが回答を行った.

### 2. 第一部: 情報提供

まず江間有沙が人工知能と社会に関する研究・政策を紹介した. 現在, 人工知能開発に10年で1,000億円を投じるといわれているが, 1980年代にも第五世代コンピュータなどの大型プロジェクトが展開されていた. 同時期に情報倫理の構築も行われており, ACMやIEEEなど, 現在も広く認知されている倫理綱領を1990年代初めに制定した. 日本においても1996年に情報処理学会, 1998年に電子情報通信学会が倫理綱領を策定した.

現在の第3次人工知能ブームでは技術開発と同時に制度や倫理も考えることが必須となっている. 例えば第5期科学技術基本計画では, 学会としての倫理ガイドラインの策定などが望まれている [内閣府 16]. 国も制度的枠組みの構築に意欲的であり, 総務省情報通信政策研究所のAIネットワーク化検討会議では, 研究開発に関する原則や解決すべき課題などのリストが列挙されている [総務省情報通信政策研究所 16].

倫理委員会は活動を開始して以来, 声明を出すことを検討してきた [松尾 15]. 倫理綱領は専門家集団に求める倫理規範の原理原則を定めた声明である. 一方, ア

メリカ人工知能学会 (AAAI) のパネル [AAAI 09] や Future of Life Institute のオープンレター [Future of Life Institute 15] のように, 人工知能の長期的・短期的なリスクやその倫理的・法的な課題について一般に向けて広く議論を呼びかけるような声明を出す方法もあり, その内容が紹介された.

続いて, 委員長の松尾 豊氏が倫理委員会の活動を報告した. 倫理委員会はこれまで人工知能が社会に与える影響を議論してきた. しかし, 研究者間でも人工知能の定義は明確ではなく, 「人工知能とは」を議論するだけで一冊の本ができてしまうくらいである [人工知能学会 16]. 一方, 世の中は人工知能ブームとなっており, 江間が紹介したように, 国内外でさまざまな議論が展開されている. 極端な例だが, 明日ニュースで「人工知能の三原則が発表」などと報道されてもおかしくはない. そのような中, 人工知能を深く理解をしている学会が倫理の問題を議論し, 社会に発信していくことが重要である. その一つの形として今回, 倫理綱領を策定したと説明した.

次に, 塩野 誠氏が今回の倫理綱領案の成り立ちと構成を説明した. 倫理綱領案は, 倫理委員会で培われた議論のほか, 他の科学技術系学会における倫理綱領などを参考として作成された. 序文では, 人工知能が将来的に自律性をもつ脅威があることを, 研究者が自覚したうえで倫理的に行動すべきであると記している. 綱領案は

1. 人類への貢献, 2. 誠実な振舞い, 3. 公正性,
4. 不断の自己研さん, 5. 検証と警鐘,
6. 社会の啓蒙, 7. 法規制の遵守, 8. 他者の尊重,
9. 他者のプライバシーの尊重, 10. 説明責任

の10条項からなる.

綱領案に対し, 情報処理学会, 電子情報通信学会の倫理綱領策定にも携わった経験をもつ土屋がコメントした. 人工知能が注目されている現在, 学会としての対応が必要であり, そのために倫理綱領を制定したいという意図について評価したい, と土屋氏は述べた.

人工知能研究者は人工知能が存在している社会の一要素に過ぎない. 人工知能と社会の関係やリスクを議論する前に, まずは専門家集団としての信頼を得ることが大事である. そのため, 倫理綱領では研究者が「市民としての」倫理と統合的な倫理規範に従って行動していることを示したうえで, 「専門家としての」倫理規範を守るという条項を入れることが重要になるという. 前者に対

しては、草案の 8. 他者の尊重, 7. 法規制の遵守, 9. 他者のプライバシーの尊重, 3. 公正性が相当する. 後者は, 2. 誠実な振舞い, 5. 検証と警鐘, 6. 社会の啓蒙, 10. 説明責任, 4. 不断の自己研さんが相当する.

現状では内容の重複が見られるため整理が必要であるが, それに加えて「これらの条項を遵守させる」というメタ条項が必要であるとの指摘があった. また, 人工知能技術ならではの特徴を盛り込むこと, 「できないことを宣言していないか」など表現の確認が必要だと指摘した. さらに倫理綱領で専門家としての立場を表明したのちに, 具体的な行動指針や事例集の策定が望まれるとのコメントで発表を締めくくった.

### 3. 第二部：公開討論

第二部は武田英明氏の司会のもと, 土屋氏と倫理委員会のメンバを合わせた 8 名で議論を行った. 土屋氏の「できないことを宣言していないか」との問題提起を受け, 学会員の服部宏充氏や西田豊明氏が, 研究者としてどこまで自分のつくったものに責任をもてるのか, 例えば綱領案の第 10 項に掲げられている「説明責任」を研究者はどこまでできるのかという問題を提示した. 松尾氏も, 学習する人工知能は将来的には自律性をもつことが期待され, 想定外の振舞いをする可能性があることは確かであるが, そもそも人工知能の研究者は人間の知能の崇高さや深遠さを解明したいと考えており, それを目指している気持ちは倫理綱領で表現できないだろうかと述べた. それを受けて土屋氏は, すべてのリスクを事前に列挙することは不可能であり, どこかで責任の線引きをする必要があると指摘した. 機械は複雑であり副作用や想定外の振舞いをすることもある. その中でより「良い」ものをつくるために, 研究者がどのように技術や社会と関わるかの議論が大事となる. また, 学会は研究者の行動を規制するだけでなく研究の振興も行う. そのバランスをどうとるかも考えていってほしいと期待を述べた.

自律的な人工知能の脅威については, 軍事にも使えてしまうデュアルユーステクノロジーを倫理綱領ではどう扱うのかとの質問が会場から寄せられた. それに対して土屋氏は, 「研究成果を論文で公開すべき」といった条項を倫理綱領に盛り込めば, 軍事研究という表現をせずとも事実的には軍事研究をしにくくなるのでは, と提案した.

このように人類を滅ぼす人工知能の議論が展開されていた中, 雇用の問題や気付かないうちに技術に我々がコントロールされているのではないかなど, ほかに重要な問題があるのではないかと服部氏が指摘した. 社会的弱者が不利な立場に置かれる事態となった場合, それに対して研究者はどのような態度をとるべきか. 綱領案の第 8 項には「他者の尊重」が掲げられており, 塩野氏は知らないうちに人工知能をもつ者, もたざる者によって

支配や搾取が起きてはならないと述べた. また, そのような漠然とした不安があることに対して研究者が答えていくべきである, と専門家としての責任について考えを述べた.

長谷敏司氏は, 人工知能研究者ではない立場からすると, 研究者にはもっと発言をしてほしいと要望を伝えた. 人工知能はそれ自体が判断能力をもっており, 大量の情報を整理するため政治性を帯びる. そのため, 研究している人は誰なのか, また, 何を考えているのかについて積極的に公開してほしいと述べた. また江間も人工知能はビッグデータと切り離すことはできず, すでに社会的なデータのみならず医療データなどが収集され管理されているということ, 我々自身が自覚することも大事ではないかと述べた.

議論の後半では, 倫理綱領という形式は本学会が提示するものとして適切かどうかの議論が行われた. 西田氏は, 倫理綱領は研究者の信頼性を担保することを意図した職業倫理に由来するが, いま人工知能研究者に求められているものは, 自分達が実現しようとしている人工知能が人間社会に災いをもたらしものであってはならないといった考えを宣言し, 人工知能の特徴を考慮に入れた倫理がどのようなものか, 考えを示すことではないだろうかと問題を提起した.

司会の武田も, 倫理綱領は学会の方向性や理念を示せる一方で, 同意できなければ行動や研究に制約を課するというデメリットがあるのでは, と述べた. それに対し土屋氏は反論する. 人工知能技術が一般の生活で利用されている現状を鑑みたとき, 人工知能そのものの倫理を議論する前に, 研究者自身がどのように社会に関与していくのかという立場を定めることが重要であり, その表現手段が倫理綱領である. また, 綱領に従えないから即除名とせずとも, 理由を述べてもらうなどの柔軟な綱領の在り方を提示した. 塩野氏も, 国内外の外部機関が制定した開発原則などによって研究に制限をかけられる前に, 専門家自らが綱領を出すことが重要ではないかと述べた. 特に研究だけではなくビジネスの現場でも, 参照する規範が求められていると指摘した. 長谷氏も, 研究者が一枚岩ではないということに理解を示しつつも, 学会員は綱領に従っているとわかることで, 一般の人々とコミュニケーションが取りやすくなるのではないかと指摘した. 長倉克枝も記者としての立場から, 倫理綱領があることによって研究者の信頼性担保につながるのではないかと述べた.

最後に司会の武田が, 「倫理綱領が必要だと思うか」について登壇者と会場で採決を行ったところ, 多くは倫理綱領が必要であると手をあげた. これに対し, 西田氏は専門家に対する倫理綱領が必要かと問われれば賛成はするが, 「人工知能の倫理について社会と一緒に考えるための宣言を策定する」などの選択肢があるならば, そちらを自身は優先したいと述べた. 松尾氏は, 人工知能



図1 公開討論の様子

研究者が社会に対して責任を果たしたいと考えていることを表現して発信していくためにも、倫理綱領のほかに人工知能に対する深い洞察に基づいた議論を行っていくことが重要であると述べた。これを受けて長倉克枝や江間が、ホームページなどのプラットフォームを使い、倫理委員会の活動としても双方向で対話をしていくような体制を整えていきたいと述べた。

公開討論会では倫理綱領案を材料として専門家として

の責任や社会との対話の在り方、また人工知能の倫理を広く社会とともに考えていくような人工知能宣言のようなものの必要性について議論が行われた。今後も倫理委員会の活動を継続していくと同時に、ホームページなどで情報を発信し対話を行っていききたい。

#### ◇ 参考文献 ◇

- [AAAI 09] AAAI Presidential panel on long-term AI futures: 2008-2009 study : <http://www.aaai.org/Organization/presidential-panel.php>
- [Future of Life Institute 15] Future of Life Institute : <http://futureoflife.org/>
- [人工知能学会 16] 人工知能学会 監修：人工知能とは，近代科学社（2016）
- [人工知能学会倫理委員会 16] 人工知能学会 倫理綱領案 : [http://ai-elsi.org/wp-content/uploads/2016/06/倫理綱領案\\_Ver3.1.pdf](http://ai-elsi.org/wp-content/uploads/2016/06/倫理綱領案_Ver3.1.pdf)
- [松尾 15] 松尾 豊，西田 豊明，堀 浩一，武田英明，長谷敏司，塩野 誠，服部宏充：ア－ティクル：人工知能学会倫理委員会の取組み，人工知能，Vol. 30, No. 3, pp. 358-364（2015）
- [内閣府 16] 内閣府：第5期科学技術基本計画（2016）：<http://www8.cao.go.jp/cstp/kihonkeikaku/index5.html>
- [総務省情報通信政策研究所 16] 総務省情報通信政策研究所 AI ネットワーク化検討会議報告書 2016 : [http://www.soumu.go.jp/main\\_sosiki/kenkyu/iict/](http://www.soumu.go.jp/main_sosiki/kenkyu/iict/)