

人工知能学会倫理委員会を振り返る

武田 英明

国立情報学研究所

takeda@nii.ac.jp

人工知能学会 倫理委員会

- 委員：
 - 松尾豊(委員長、東大)
 - 西田豊明(京大)
 - 堀浩一(東大)
 - 武田英明(NII)
 - 服部智充(立命館)
 - 長谷敏司(SF作家)
 - 塩野誠((株)経営共創基盤)
 - 江間 有沙(東大)
 - 長倉 克枝(科学ライター)
- オブザーバ
 - 松原仁(公立はこだて未来大学)
 - 山川宏(ドワンゴ人工知能研究所)
 - 栗原聡(慶應義塾大学)

人工知能学会 倫理委員会

- 2014年9月発足
- 2014年12月第1回委員会
- 2015年2月第2回委員会
- 2015年6月第3回委員会(全国大会での公開委員会)
- 2015年10月第4回委員会
- 2016年3月第5回委員会
- 2016年3月第6回委員会
- 2016年6月第7回委員会
- 2016年6月全国大会での公開討論会(NHK,日経、朝日等報道)
- 2016年9月第9回委員会
- 2016年12月第10回委員会
- 2016年2月第11回委員会
- 2017年5月全国大会で公開討論会を実施
- 2017年7月第12回委員会
- 2017年10月第13回委員会
- (2018年6月全国大会で企画セッション開催
「AIに関わる安全保障技術を巡る世界の潮流」)

「AIに関わる安全保障技術を巡る世界の潮流」

2018年6月5日 17:20-19:00
人工知能学会全国大会
鹿児島市

- 安全保障技術の国際的な動向について(拓殖大学国際学部教授・海外事情研究所副所長 佐藤丙午氏)
- LAWSと国連の動向などについて(外務省軍縮不拡散・科学部通常兵器室上席専門官の南健太郎氏)



NEWS

[RSS](#)

2017年02月28日 お知らせ
「人工知能学会 倫理指針」が、理事会で承認されました

2016年06月30日 お知らせ
2016年度人工知能学会全国大会「公開討論：人工知能学会 倫理委員会」の開催報告および、会場で募集したコメントへの回答を掲載しました

2016年06月09日 お知らせ
2016年度人工知能学会全国大会「公開討論：人工知能学会 倫理委員会」での発表資料を掲載しました

2016年06月07日 お知らせ
人工知能研究者の倫理綱領(案) を公開しました

2016年06月06日 お知らせ
人工知能学会倫理委員会Webサイトを公開しました

ニュース

動画

特集

スペシャルコンテンツ

NEWS WEB EASY

新着

社会

気象・災害

科学・文化

政治

ビジネス

国際

スポーツ

暮らし

地域

宮城福島震度5弱

ジョンナム氏殺害

大阪 国有地売却

トランプ大統領

働き方改革

豊洲市場問題

▶ 注目ワード一覧を見る

News Up

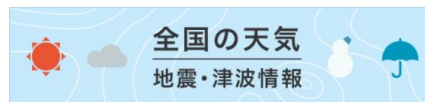


人工知能 研究者が守るべき倫理指針 学会がまとめる

2月28日 20時33分

急速に開発が進んでいる人工知能が人類への脅威とならないよう、研究者が守るべきルールを定めた初の倫理指針を国内4000人の研究者でつくる人工知能学会がまとめました。他人に危害を加える意図をもって利用することを禁じるとともに、人工知能自体も倫理を守らなければならないと強調しています。

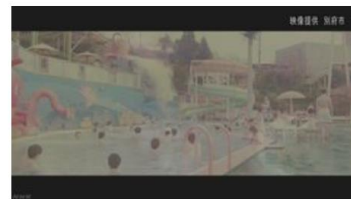
人工知能をめぐるのは、人間のようにみずから考えて行動する“自律性”を目指して急速に研究開発が進む一方、人工知能を悪用したり、将来、人工知能が人の指示に反したりすることも懸念されています。



気象

雨雲データマップ

ソーシャルランキング



1

別府の「湯〜園地」実現へ ネットで募った資金が目標額に

2

宇宙ごみ対策 “衛星が燃え尽きる” 世界初の技術実験に成功

3

卒業式終えた高校生が白線流し 岐阜 高山

4

90歳男性が高速道路逆走 大型トラックと正面衝突し死亡

5

自民 鴻池元防災担当 3年前に森友学園理事長夫妻と面会

人工知能学会 倫理指針(その1)

序文

人工知能研究は、人間のような知性を持ち自律的に学習し行動する人工知能の実現を目指している。人工知能が、産業、医療、教育、文化、経済、政治、行政など幅広い領域で人間社会に深く浸透することで、人々の生活が格段に豊かになることが期待される一方で、悪用や濫用で公共の利益を損なう可能性も否定できない。

高度な専門的職業に従事する者として、人工知能の研究、設計、開発、運用、教育に広く携わる人工知能研究者は、人工知能が人間社会にとって有益なものとなるようにするために最大限の努力をし、自らの良心と良識に従って倫理的に行動しなければならない。人工知能研究者は、社会の様々な声に耳を傾け、社会から謙虚に学ばなければならない。人工知能研究者は技術の進化及び社会の変化に伴い、人工知能研究者自身の倫理観を発展させ深めることについて不断の努力をおこなう。

人工知能学会は、自らの社会における責任を自覚し、社会と対話するために、人工知能学会会員の倫理的な価値判断の基礎となる倫理指針をここに定める。学会員はこれを指針として行動するよう心がける。

人工知能学会 倫理指針(その2)

1. **(人類への貢献)**人工知能学会会員は、人類の平和、安全、福祉、公共の利益に貢献し、基本的人権と尊厳を守り、文化の多様性を尊重する。人工知能学会会員は人工知能を設計、開発、運用する際には専門家として人類の安全への脅威を排除するように努める。
2. **(法規制の遵守)**人工知能学会会員は専門家として、研究開発に関わる法規制、知的財産、他者との契約や合意を尊重しなければならない。人工知能学会会員は他者の情報や財産の侵害や損失といった危害を加えてはならず、直接的のみならず間接的にも他者に危害を加えるような意図をもって人工知能を利用しない。
3. **(他者のプライバシーの尊重)**人工知能学会会員は、人工知能の利用および開発において、他者のプライバシーを尊重し、関連する法規に則って個人情報 の適正な取扱いを行う義務を負う。
4. **(公正性)**人工知能学会会員は、人工知能の開発と利用において常に公正さを持ち、人工知能が人間社会において不公平や格差をもたらす可能性があることを認識し、開発にあたって差別を行わないよう留意する。人工知能学会会員は人類が公平、平等に人工知能を利用できるように努める。
5. **(安全性)**人工知能学会会員は専門家として、人工知能の安全性及びその制御における責任を認識し、人工知能の開発と利用において常に安全性と制御可能性、必要とされる機密性について留意し、同時に人工知能を利用する者に対し適切な情報提供と注意喚起を行うように努める。

人工知能学会 倫理指針(その3)

- 6 (誠実な振る舞い)人工知能学会会員は、人工知能が社会へ与える影響が大きいことを認識し、社会に対して誠実に信頼されるように振る舞う。人工知能学会会員は専門家として虚偽や不明瞭な主張を行わず、研究開発を行った人工知能の技術的限界や問題点について科学的に真摯に説明を行う。
- 7 (社会に対する責任)人工知能学会会員は、研究開発を行った人工知能がもたらす結果について検証し、潜在的な危険性については社会に対して警鐘を鳴らさなければならない。人工知能学会会員は意図に反して研究開発が他者に危害を加える用途に利用される可能性があることを認識し、悪用されることを防止する措置を講じるように努める。また、同時に人工知能が悪用されることを発見した者や告発した者が不利益を被るようなことがないように努める。
- 8 (社会との対話と自己研鑽)人工知能学会会員は、人工知能に関する社会的な理解が深まるよう努める。人工知能学会会員は、社会には様々な声があることを理解し、社会から真摯に学び、理解を深め、社会との不断の対話を通じて専門家として人間社会の平和と幸福に貢献することとする。人工知能学会会員は高度な専門家として絶え間ない自己研鑽に努め自己の能力の向上を行うと同時にそれを望む者を支援することとする。
- 9 (人工知能への倫理遵守の要請)人工知能が社会の構成員またはそれに準じるものとなるためには、上に定めた人工知能学会会員と同等に倫理指針を遵守できなければならない。

本指針は理事会成立後より公布する。

本指針の解釈および見直しについては、必要に応じて委員会を開催し、理事会の承認を得る。

以上

人工知能学会 倫理委員会

- 「倫理」委員会
 - 名称には議論あり。「AIと未来社会検討委員会」などの提案もあり。
 - “人工知能のプレゼンスが高まるなかで、社会との接点を真剣にかんがえないといけない”（松尾）

人工知能学会 倫理委員会

- これまでの議論項目

- 正しい現状認識の必要性

- 今すぐシンギュラリティはやってこない。
 - いまのところ、AIシステムは設計者が設計するものである。(人の関与しないAIシステムは存在しない)
 - ただし、レバレッジ効果を巨大になりつつある。(一度設計されたら、そのシステムがもたらす影響ははかりしれなく大きくなった)
 - AIシステムは人間社会にシームレスに入り込み、その影響はこれまでのコンピュータシステムより直接的に人間と社会に影響を与える

人工知能学会 倫理委員会

- これまでの議論項目
 - 我々の役割の明確化
 - 研究者の自覚を促す
 - 研究者自身に自らの研究活動が社会に大きく影響を与えう
るということを自覚させること
 - 専門家としての責任
 - 専門家として予見できるものは予見しとく
 - 社会への発信、コミュニケーション

人工知能学会 倫理委員会

- これまでの議論項目
 - 個別的トピックス
 - 設計者の倫理
 - 職業の問題
 - 人工知能の「心」の問題

人工知能設計者の倫理

- ロボット3原則はない
- 「ロボットが」守るべき原則ではなくて、「ロボット設計者が」守るべき原則が必要

人工知能設計者の倫理

- 生命科学では一定の倫理規定がある。
- 人工知能でも同様の倫理規定が作れるのではないか。

人工知能設計者の倫理

- 人工知能設計者が守るべきことはあるのか。
- Q: 人を殺すAIは設計してはいけないのか？
(自動運転車は絶対に人を殺さないとはいえない)
- Q: AIが人の命の軽重を判断するようにして良いのか(避けられない状態での選択)
- Q: AIシステムが殺人などや法を反することをしたら、それは設計者の責任か
- ...

人工知能設計者の倫理

- 武田の答え：“同様”の倫理規定は無理
- 倫理規定に関する議論は、現在の科学技術水準による行動が社会にどのような影響を与えるかが予見できるときは、建設的議論になるであろう
 - 例えば人の卵子に対する操作がもたらす社会的インパクトを想定できるからこそ、それに対する倫理が議論できる
- AIシステムは社会に入り込み社会を変える。変わったあとの社会でしか評価できないとするなら、いま議論したことはどういう意味があるだろう。