

2016/6/6

人工知能学会倫理委員会

### 注意事項

本倫理綱領案は2016年度人工知能学会全国大会における倫理委員会企画による公開討論での意見募集のために作成された現時点での案文であり、人工知能学会の公式の見解や意向を示す最終化されたものではないことをご了解ください。

## 人工知能研究者の倫理綱領(案)

### 序文

人工知能は経済、政治、産業、行政、教育、医療といった幅広い領域で人類にとって目に見える形と見えない形で大きな役割を果たしていくものと考えられる。人工知能研究開発者は人工知能の研究、設計、開発、運用、そして研究者の教育を行っている。人工知能はその汎用性と潜在的な自律性から人工知能研究開発者の想定しえない領域においても人類に影響を与える可能性があり、人工知能研究開発者によって為された研究開発がその意図の有無に関わらず人間社会や公共の利益にとって有害なものとなる可能性もある。

人工知能の研究開発が人間社会にとって有益なものとなるために、人工知能研究開発者は高度な専門的職業に従事する者として自らの良心と良識に従って倫理的に行動すべきである。人間が創造したものによって人間の幸福を損ねることがあってはならないことを人工知能研究開発者はここに確認し行動する。

倫理綱領はその総体によって人工知能研究開発者の倫理的な価値判断の基礎となるものであり、その一部のみを捉えて人工知能研究開発者の行為を正当化するものではない。人工知能研究開発者は専門的知識を持たない者に対しても敬意と想像力を持って接し、人工知能研究開発者の行為に対する社会的な評価に絶えず注意を払うべきである。

本倫理綱領は人工知能研究開発者のすべてに適用されるべきであり、人工知能研究開発者はここに定められる倫理綱領を遵守することを誓うものとする。

### 綱領

1 (人類への貢献) 人工知能研究開発者は人類の平和、安全、福祉、公共の利益に貢献

し、基本的人権を守り、文化の多様性を尊重する。人工知能研究開発者は人工知能を設計、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しなくてはならない。

- 2 **（誠実な振る舞い）** 人工知能研究開発者は社会に対して誠実に信頼されるように振る舞う。人工知能研究開発者は専門家として虚偽や不明瞭な主張を行わず、人工知能を構成するシステムの技術的限界や問題点について科学的に真摯に説明する義務を負う。
- 3 **（公正性）** 人工知能研究開発者は常に公正さを持ち差別を行わない。人工知能の利用が人間社会において新たな不公平や格差をもたらす可能性があることを認識し、人種、国籍、宗教、性別、職業、障がい等による差別を行わない。人工知能研究開発者は人類が公平、平等に人工知能という資源を利用できるよう最善を尽くすものとする。
- 4 **（不断の自己研鑽）** 人工知能研究開発者は高度な専門家として絶え間ない自己研鑽に努め、その能力の向上を行う。
- 5 **（検証と警鐘）** 人工知能研究開発者はその研究開発し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。
- 6 **（社会の啓蒙）** 人工知能研究開発者は人工知能に関する社会的な理解を向上させるよう努力しなければならない。人工知能研究開発者は自らの良心と良識に従って人工知能の技術的可能性と限界について社会全般を啓蒙すると同時に、社会が人工知能に対して誤った認識を有している場合にはそれに対し自らが正しいと信じる主張を行うべきである。
- 7 **（法規制の遵守）** 人工知能研究開発者は専門家として研究開発に関わる法規制を遵守し、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。現行の法規制が技術に整合していないと人工知能研究開発者が自ら倫理的に判断する場合には、専門家としての行為と結果に対し責任を持たなければならない。
- 8 **（他者の尊重）** 人工知能研究開発者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ない。人工知能研究開発者は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てはならない。人工知能研究開発者は許可なく他者のシステムやデータにアクセスを行わない。こうしたアクセスも他者への危害と同等に扱われる。人工知能研究開発者は他者に危害を加える可能性を最小限にし、意図せざる危害を他者に及ぼしてしまった際にはその損害を回復し、回復が困難

な場合には損害を緩和する措置を講じるべきである。

- 9 **（他者のプライバシーの尊重）** 人工知能研究開発者と人工知能を内包したシステムは他者のプライバシーを尊重する。予め定められた目的のために集められた個人情報は、本人の同意なくしては他の目的のために利用されず、人工知能研究開発者は個人情報の適正な取扱いを行う義務を負う。人工知能の運用において得られた個人情報は濫用されることなく内密に扱われなければならない。
  
- 10 **（説明責任）** 人工知能研究開発者は研究開発が意図せず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、もしも人工知能が悪用されていることを発見した際には、技術を悪用する者に対して説明を求め、その説明が正当なものでない場合には、悪用されることを防止する措置を講じなければならない。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はあってはならない。

以上