

## S-HIGO フェローシッププログラムにおける対話型 AI の取扱いについて

令和 5年 5月 31日 運営委員会決定

OpenAI 社が公開している ChatGPT などの対話型 AI は、情報収集や整理などの機械的な作業を自動的に行うことで知的作業の飛躍的な効率化をもたらすという利点があります。また、AI との対話を通してこれまで気付かなかつた視点に気付き、新しい着想に至るという効果もあります。対話型 AI を含む生成 AI は、インターネットのようにこれからの社会に欠かすことのできない技術になると考えられます。

対話型 AI は入力したプロンプトに対して尤もらしく説得力を感じる回答を生成するので、レポートや論文などの科学的論述をそれらしく簡単に作成することができます。しかし、対話型 AI はプロンプトと学習したデータセットの統計的関連性に基づいて文章を生成するため、生成された回答には誤った情報が含まれていたり、情報にバイアスがかかっていたりします。誤った情報やバイアスのある情報を鵜呑みにすることによって、正しい論述ができないだけでなく、その情報を基にして自分の研究を進めた場合、誤った方向に導かれる危険性もあります。したがって、対話型 AI を正しく利用するためには、生成された回答が科学的に正しいかどうかを批判的に判断できる専門的な知識が必要になります。そのような知識を持った上で対話型 AI を活用すれば作業の効率化も可能になります。

一方、特定の課題について論述する能力が自分にないため、対話型 AI の回答をそのまま(若しくは少し表現を変えて)無批判に使用すると、誤った情報を含む論述を提示することになりかねません。そもそも、他者の論述を自分の論述として提示することは、既存の論文やインターネット上の情報をコピーペーストする場合と等しく剽窃行為にあたります。それでは、対話型 AI、例えば「ChatGPT」を引用元として表示すれば良いのでしょうか。科学的な論述で最も大切なことは、論述する者はその論述の根拠を示す説明責任を負うということです。対話型 AI は現在のところ引用元を正しく示すことができず、論述の正しさを保証できず、論述に対する説明責任を果たせないため、一般的に論文の(共)著者としては認められません。著者として認められない対話型 AI を引用元として表示することは適切ではありません。

また、対話型 AI には透明性についての懸念もあります。例えば、ChatGPT が使用する大規模言語モデルや学習に利用されるデータセットは公開されていません。プロンプトに入力した情報は他の文章の生成に引用無く使用される可能性があり、個人情報や大切な研究情報の漏洩につながる危険性も指摘されています。

これまで述べたように、対話型 AI を科学的な論述に利用する場合、常に真実性、誠実性、透明性を意識する必要があります。そこで S-HIGO フェローシッププログラムでは、レポート、研究計画書・報告書、ポスター等の作成に対話型 AI を利用したかどうか、利用した場合はどの対話型 AI をどのように利用したかを明示することを求めることとします。

科学研究における知的営みの全てを AI に頼るとすれば、人間の研究者としての自律性は失われます。しかし、将来いくら AI 技術が発達しても、人間自身の好奇心、想像力、創造性、独創性がなければ、科学の進展もまたあり得ないでしょう。AI 技術が急速な進化を遂げている今、AI 技術にどのように向き合い、どのように正しく科学研究に役立てるのかをしっかりと考えることがますます大切になっています。

(参考)

Nature editorials 「Tools such as ChatGPT threaten transparent science; here are our ground rules for their use」 Nature 613, 612 (2023)

<https://doi.org/10.1038/d41586-023-00191-1>

E. A. M. van Dis *et al.* 「ChatGPT: five priorities for research」 Nature 614, 224–226 (2023)

<https://doi.org/10.1038/d41586-023-00288-7>

H. Yang 「How I use ChatGPT responsibly in my teaching」 Nature (Career Column, 12 April 2023)

<https://doi.org/10.1038/d41586-023-01026-9>

太田邦史 「生成系 AI (ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion 等) について」 (2023 年 4 月 3 日) 東京大学 utelecon: オンライン授業・Web 会議ポータルサイト

<https://utelecon.adm.u-tokyo.ac.jp/docs/20230403-generative-ai>

(注)この文書の作成にあたって対話型 AI は使用されていません。