AI(人工知能)(5)

2024年2月20日作成

三木

◆学習データの所在

生成AIは大量の学習データを必要とする。

大量の学習データを取得するには、インターネット上のデータを利用するのが効率的である。

そうであれば、インターネット上のデータでなければ、生成AIに学習されない。

◆学習データの信頼性

インターネット上には大量のデータがある。

しかし、それらは玉石混合であるため、生成AIに無秩序に学習させるとは考えにくい。

おそらく、信頼性の高いデータを優先的に学習させるだろう。

信頼性の高いデータとは、学術論文・解説、百科事典、新聞記事、書籍などである。

そうであれば、インターネット上のデータであったとしても、信頼性が高くなければ、生成AIに学習されない。

◆当会の知見

当会の知見もできれば生成AIに学習させたい。つまり、生成AIに影響を与えたい。そうしなければ、生成AIの世界において、当会の知見は存在しないことになる。

当会の知見は、書籍やHPの資料で公開されている。

しかし、書籍はインターネット上のデータではないため、このままでは生成AIに学習されない。

しかるべき時期や方法を検討して、インターネット上に公開すべきであろう。

また、会員限定とされている資料も、適切に公開すべきであろう。

◆当会の信頼性

当会は非営利活動法人であるため、少なくとも不特定多数の個人に比べると信頼性は高いはずである。しかし、著名な団体や個人に比べると信頼性は劣る。

そのため、信頼性の高いデータから当会への誘導を検討すべきであろう。例えば、代表的な学習データとされるWikipediaに当会の項目を加えることも一つの方法である。

以上