

新たな情報財検討委員会報告書 抜粋

(平成29年3月 知的財産戦略本部 検証・評価・企画委員会 新たな情報財検討委員会)

第2. AIの作成・利活用促進のための知財制度の在り方

2. 論点

(1) AI学習用データに関する論点(「データ作成者」と「AI学習を行う者」が異なる場合の著作権法上の課題等)

＜現行知財制度とAI学習用データ作成に係る状況＞

学習用データを作成するにあたり、元となるデータに著作物が含まれている場合¹であっても、著作権法第47条の7²に基づき、必要な限度で著作物を記録又は翻案し、学習用データを作成することが可能と考えられる³。一方で、同条では譲渡や公衆送信が規定されていない。

このため、複数の当事者が協業で学習済みモデルを作成しようとした場合に、学習用データを作成する⁴主体(以下「データ作成者」という。)と、実際にAI学習⁵を行う主体(以下「AI学習を行う者」という。)が異なるとき、データ作成者からAI学習を行う者へ学習用データを提供又は提示する行為⁶が著作権法上違法と解されるおそれがある。

＜検討の視点＞

AIの作成を促進する観点から、学習用データの作成や提供・提示を促進することが求められ、これに向けた課題についてどのように考えるかが検討の視点となる。

課題①-1) AI学習用データの作成・特定当事者間の提供及び提示

現在のAIの作成・開発においては、分析等に用いる学習用データを収集するデータ作成者とAIの開発に関する技術を有しているAI学習を行う者が異なる場合が多く、これらの者が協業することによりビジネスとしてAIの開発を行っているとの指摘がある。このような場合、データ作成者からAI学習を行う者に対して学習用データが提供又は提示されることは一般的に行われていると考えられる。

しかし、このような場合に、学習用データをAI学習を行う者へ提供又は提示する行為が著作権法上違法と解されてしまうと、協業によるAIの作成・開発に支障

¹ 例えば、インターネット上のデータを元に学習用データを作成する場合が考えられる。

² 著作権法(抜粋)

第47条の7 著作物は、電子計算機による情報解析(多数の著作物その他の大量の情報から、当該情報を構成する言語、音、映像その他の要素に係る情報を抽出し、比較、分類その他の統計的な解析を行うことをいう。以下この条において同じ。)を行うことを目的とする場合には、必要と認められる限度において、記録媒体への記録又は翻案(これにより創作した二次的著作物の記録を含む。)を行うことができる。ただし、情報解析を行う者の用に供するために作成されたデータベースの著作物については、この限りでない。

³ なお、同条但書により、情報解析を行う者の用に供するために作成された「データベースの著作物」は適用除外となっている。

⁴ 「記録媒体へ記録又は翻案を行う」に該当すると考えられる。

⁵ 「電子計算機による情報解析」に該当すると考えられる。

⁶ 学習用データの記録媒体を譲渡する方法で提供またはクラウド等に学習用データをアップする方法などが考えられる。

が生ずるおそれがある。

これについては、「公衆」への提供や提示に該当しない特定当事者間における記録媒体の譲渡又は特定当事者にアクセスを制限するような形式でクラウド上へアップするような送信行為であれば現行法上も違法ではないとの指摘がある一方⁷で、著作権法上の「特定」は、裁判例⁸や学説において個人的な結合関係にあるものを指すとされており、かなり限定的に解釈されるとの指摘もあった。

この点、A I の作成・開発の促進という観点からは、情報解析という共通の目的下での「公衆」に該当しない特定当事者間での提供・提示であれば、現行の著作権法でも可能と解釈されるべきであると考えられる⁹。

課題①－2) 民間におけるA I 学習用データの作成と特定当事者間を超えた提供・提示

A I の研究・開発を活性化させるため、著作権法第 47 条の 7 又は同様の権利制限規定により、作成した学習用データについて、特定当事者間を超えて、現行著作権法で認められていない公衆への提供を可能にすべきではないかとの指摘がある¹⁰。

具体的に、特定当事者間であるとは言えない可能性のある多数の者が参加するコミュニティにおいて、学習用データ等を共有してA I の研究・開発を実施した方が学習効率が向上するとの指摘があった¹¹。

そもそも、A I の性能は、学習用データの量や内容等に左右されるとの指摘があり、多数の者が共同でデータを収集することでA I の性能を高めることができると考えられる。特に、現在、A I の研究・開発のために共有されている学習用データは、海外において生成されたものが多く、我が国におけるA I の研究・開発にとって十分なものとは言えない状況であるとの指摘があるため、今後、我が国の社会から生成される学習用データが、可能な限り、共同で効率よく収集され、A I の研究・開発のために共有されることが望ましいと考えられる。

しかし、学習用データの公衆への提供を可能としてしまうと、「学習用データ」と称して著作物の表現を知覚することが容易なデータが頒布されるおそれがあり、権利者への影響が大きいと考えられる。

この点を含む次世代知財システムの在り方については、次世代知財システム検討

⁷ 文化審議会著作権分科会法制・基本問題小委員会中間まとめ（平成 29 年 2 月）43 頁注 61 にも現行法の解釈によって対応可能との意見が記載されている。

⁸ NTT リース事件（東京地裁平成 16 年 6 月 18 日）

原告が著作権を有するプログラムについて、被告が使用許諾契約で定めた再使用許諾の範囲を超えて少数のグループ会社に再使用許諾をした事案において、『特定』というのは・・・人的な結合関係が存在することを意味する」と述べた上で、リース会社にとってのリース先は人的な結合関係はないことから「特定」とはいえず「公衆」に該当するとし、少数のグループ会社であっても「公衆」にあたりと判断した。

⁹ 平成 29 年 2 月に公表された文化審議会著作権分科会法制・基本問題小委員会中間まとめにおいて、柔軟性の高い規定を整備することが望ましいとされている「著作物の本来的利用には該当せず、権利者の利益を通常害さないと評価できる行為類型[第 1 層]」の中で、「技術の進展に伴い、現行規定に定める利用行為に類するものであるものの現行規定の対象範囲から外れるおそれのある行為が新たに生じてきている」例として、著作権法第 47 条の 7 について、複数の主体が協業で情報解析用データベースの作成と情報解析を分担して行うことは、権利者の利益を害するものでないことから権利制限の対象となるべき行為である旨の意見があったことが示されている（文化審議会著作権分科会法制・基本問題小委員会中間まとめ（平成 29 年 2 月）43 頁注 61）。

¹⁰ 米国ではインターネット上の画像を学習用データに加工して公衆に向けて提供しているサイト（例：ImageNet (<http://image-net.org/index>)) があるとの指摘があったが、このようなサイトについては適法であるとの判決があるわけではなく、合法であるか明確でないとの指摘があった。

¹¹ 新たな情報財検討委員会第 3 回資料 4（産総研提出資料）参照

委員会報告書（平成28年4月）において「人工知能が学習等するために必要なビッグデータの収集・蓄積と利活用が行われやすい環境整備が重要」とされたことを踏まえ、知的財産推進計画2016においては、デジタル・ネットワーク時代の著作物の利用への対応の必要性に鑑み、新たなイノベーションへの柔軟な対応等の観点から、著作権法の権利制限規定等について、検討を行い、必要な措置を講ずることとされている。

文化審議会著作権分科会法制・基本問題小委員会においては、著作物の表現の知覚を伴わない利用行為を含む一定の行為について柔軟性の高い権利制限規定を整備するとの方向性が示されているところであり、我が国のAIの作成の促進に向け、特定当事者間を超えて学習用データを提供・提示する行為について、新たな時代のニーズに対応した著作権法の権利制限規定に関する制度設計や運用の中で検討を進めることが適当である。

課題①-3) 公的機関等におけるAI学習用データの作成と公衆への提供

我が国のAI研究開発を活発化させるため、公衆への提供が可能なAI学習用データに関するインフラを整備すべきとの指摘がある。

これについては、官民連携で公衆への提供が可能なビッグデータのアーカイブを早急に進めるべきとの指摘や、AI学習用の図書館を作り、一定の制約の下で集めた学習用データを利用可能にすべきという指摘がある一方、利用を認める範囲について、コンテンツのようにアクセス権を有償化することでビジネスを行っている分野もあるため、フリーとする部分とそうでない部分をきちんと切り分けるべきとの指摘があった。

関係する公的機関の現場においては、各公的機関の保有するデータが自前のものか他者のデータかで法制度上の取扱いが異なることや、分野によってデータの取扱いに対する意識が異なり、統一的なルールによってオープン化を進めていくことは難しいこと等の指摘があった。

以上から、関係の公的機関がそれぞれの観点から、公開・共有することが望ましいデータとそうでないデータの整理を行った上で、公開・共有できない理由が制度的なものか、業界や分野毎に異なる共通意識等の問題であるかを整理することが求められると考えられる。まずは、業界や分野毎にデータ公開・共有への共通意識を醸成したうえで、オープンサイエンス、オープンデータ等の目的に基づき、各公的機関において適切にデータを公開・共有する取組を進めることが必要である。