

生成 AI の開発・学習段階における 著作物利用行為

弁護士・弁理士 柿沼 太一

要 約

本稿は、AI と著作権法に関する全体像について簡単に分析した後、生成 AI の開発・学習段階における著作物利用行為に焦点をあて、「学習目的による制限」と「学習対象による制限」という視点から分析したものである。「学習目的による制限」については、学習対象著作物の「表現出力目的」と「作風模倣目的」がある場合について検討し、「学習対象による制限」については、問題となることが多い「情報解析に活用できる形で整理したデータベースの著作物」「海賊版等の権利侵害複製物」「学習禁止意思が付されている著作物」「学習を防止するための機械可読方法による技術的な措置が付されている著作物」「情報解析用 DB 著作物以外の著作物・非著作物のうちライセンス市場が形成されているもの」について検討した。さらに、AI の開発・学習段階における著作権法第 30 条の 4 に関する検討結果が、AI の生成・利用段階においても基本的に当てはまることを示しつつ、法第 30 条の 4 と同じく「情報解析」に関する規定である法第 47 条の 5 と法第 30 条の 4 の役割分担についても検討した。

目 次

- 1 AI と著作権法に関する全体像
 - 1.1 分析の視点
 - 1.2 「開発・学習」段階と「生成・利用」段階の意味
 - 1.3 開発・学習段階と生成・利用段階を分けて検討する意味
- 2 生成 AI の開発・学習段階における著作物利用行為
 - 2.1 分析の視点
 - 2.2 学習目的による制限
 - (1) 表現出力目的
 - (2) 作風模倣目的
 - (3) まとめ
 - 2.3 学習対象による制限
 - (1) 情報解析に活用できる形で整理したデータベースの著作物
 - (2) 海賊版等の権利侵害複製物
 - (3) 学習禁止意思が付されている著作物
 - (4) 学習を防止するための機械可読方法による技術的な措置が付されている著作物
 - (5) 情報解析用 DB 著作物以外の著作物・非著作物のうちライセンス市場が形成されている（すでにライセンス・販売されている）もの
 - 2.4 開発・学習段階での著作権侵害行為について権利者はどの範囲で差止請求等ができるか
 - 2.5 生成・利用段階における情報解析と法第 30 条の 4
 - (1) 入力目的による制限
 - (2) 入力対象による制限
 - 2.6 法第 30 条の 4 と法第 47 条の 5 の役割分担
 - (1) 情報解析に関する著作権法上の規定
 - (2) 実際の検討の順番
 - (3) 法第 47 条の 5 は使い道があるのか

1. AI と著作権法に関する全体像

1.1 分析の視点

2024年3月15日に、文化審議会著作権分科会法制度小委員会から「AI と著作権に関する考え方について」が公表された⁽¹⁾(以下「考え方」という)。本稿では、「考え方」にも言及しながら生成 AI と著作権法に関する論点のうち、開発・学習段階における論点について検討するが、その前提として、AI と著作権法に関する論点の全体像について簡単に整理する。

AI と著作権法に関する論点の全体像については、以下の3つの視点で整理するとわかりやすい。

- ① 主体（「AI 開発者」「AI サービス提供者」「AI 利用者」）
- ② フェーズ（「開発・学習段階」「生成・利用段階」）
- ③ システム・サービスの内容（「クラウド」「ローカル」）

AI の開発から利用までを時系列順に並べると、① AI（機械学習技術を利用して作成された学習済みモデルのこと。以下同じ）。の開発・提供、②開発された AI を利用した AI サービスの開発・提供、③提供された AI サービスの利用という順になるが、本稿では、それぞれの行為を行う主体を「AI 開発者」「AI サービス提供者」「AI 利用者」とする（①主体）。もちろん、1つの主体が複数の役割（たとえば「AI 開発者」と「AI サービス提供者」）を兼ねることもある。

そして、それぞれの主体によって「AI の開発・学習」段階の行為と「AI 生成物の生成・利用」段階の行為の一方、あるいは双方が行われる（②フェーズ）。

さらに、AI サービスは、AI 利用者に対するクラウドサービスとして提供される場合もあるし、AI 利用者自身がローカル環境⁽²⁾で AI 生成物を生成・利用することもある（③システム・サービスの内容）。

最もシンプルなパターンとして「AI 開発者」と「AI 利用者」のみが登場し、「AI 開発者」は「AI の開発・学習」のみを、「AI 利用者」は「AI 生成物の生成・利用」のみを行うパターンを図示してみよう（図1）。

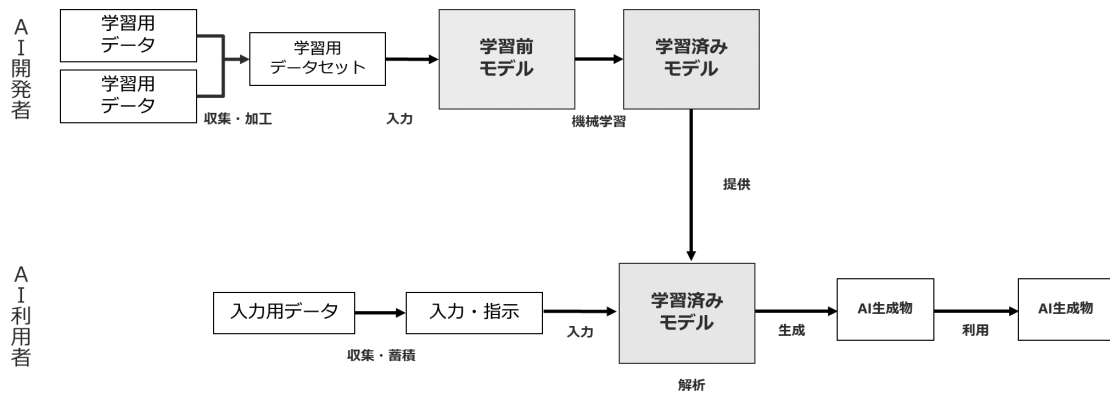


図 1

これは、「AI 開発者」が開発した AI を「AI 利用者」に提供し、AI 利用者が自らの手元で（ローカルで）当該 AI を利用して「AI 生成物の生成・利用」を行うパターンである。

次に、「AI 開発者」と「AI 利用者」のみが登場し、「AI 開発者」は「AI の開発・学習」を、「AI 利用者」は「AI の開発・学習」と「AI 生成物の生成・利用」を行うパターンを図示してみる（図2）。

(1) https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_07/pdf/94024201_01.pdf

(2) ここでいう「ローカル」には、AI 利用者自身が管理するクラウド環境も含む。

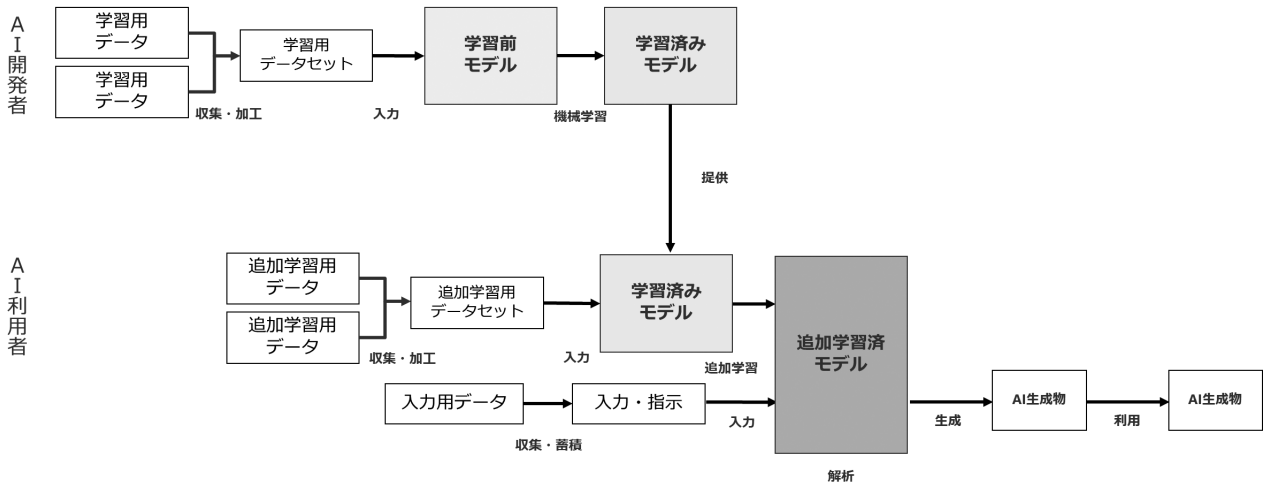


図 2

さらに「AI 開発者」「AI サービス提供者」が「AI の開発・学習」を、「AI 利用者」が「AI の開発・学習」と「AI 生成物の生成・利用」をそれぞれ行う場合は以下の図 3 のとおりとなる。

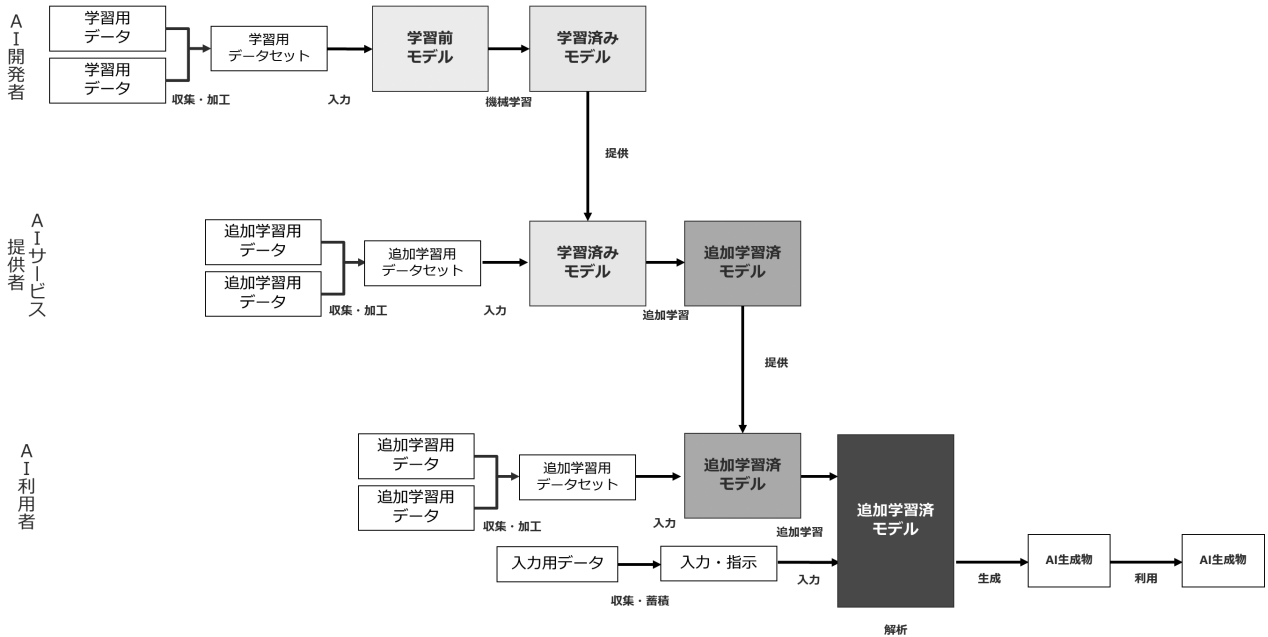


図 3

AI と著作権の問題について検討する際には、このように基本的には「主体」×「フェーズ」の組み合わせで分析することになる。

図 2 や図 3 のように、それぞれの主体において、フェーズとして「開発・学習段階」や「生成・利用段階」が複数組み合わせられることもあるが、その場合であっても、基本的に各フェーズにおける著作権侵害の有無の考え方は同一である（ただし、AI サービス提供者や AI 利用者における「学習」は少量の学習用データで学習することが多く、その際には後述の学習目的による制限がかかることもある。）。

また、図 1 ～ 図 3 では、「AI 開発者」が開発した AI を「AI 利用者」に提供し、AI 利用者が自らの手元で当該 AI を利用して「AI 生成物の生成・利用」を行うパターンであったが、実際には、AI がクラウド上のサービスとして提供されている場合もある（図 4）。

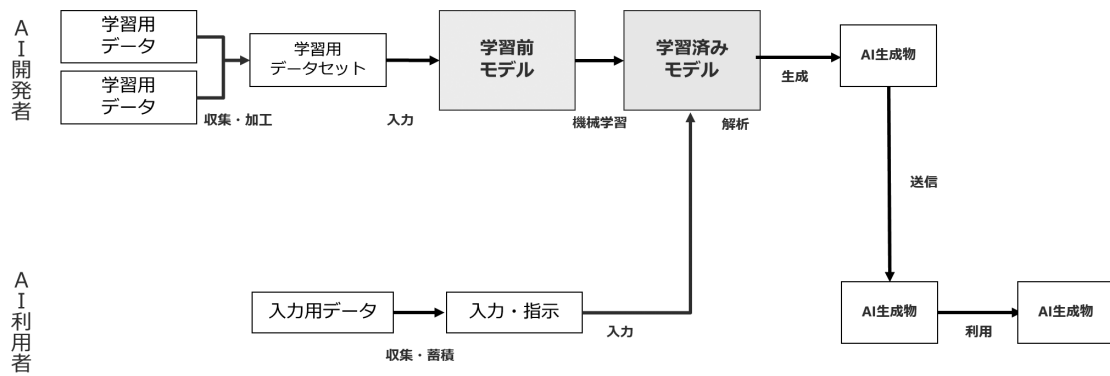


図 4

先ほどの図 1～図 3 との違いは、「生成・利用」を物理的に誰が行っているのか、という点である。

先ほどの図 1～図 3 では、「生成・利用」を物理的にやっているのは AI 利用者だが、図 4 では、物理的に見ると、「生成・利用」のうち「入力」は AI 利用者、「生成」及び生成した AI 生成物の「送信」は AI 開発者、送信された AI 生成物の「利用」は AI 利用者が行っている。

このことは、「生成・利用」について、AI 開発者がどのような責任を負うかを考えるに際して重要なポイントとなる。

1. 2 「開発・学習」段階と「生成・利用」段階の意味

ここまで特に定義することなく「開発・学習」「生成・利用」という用語を用いてきたが、本稿では、「開発・学習」段階と「生成・利用」段階という用語を、以下の図 5 の意味で用いる。

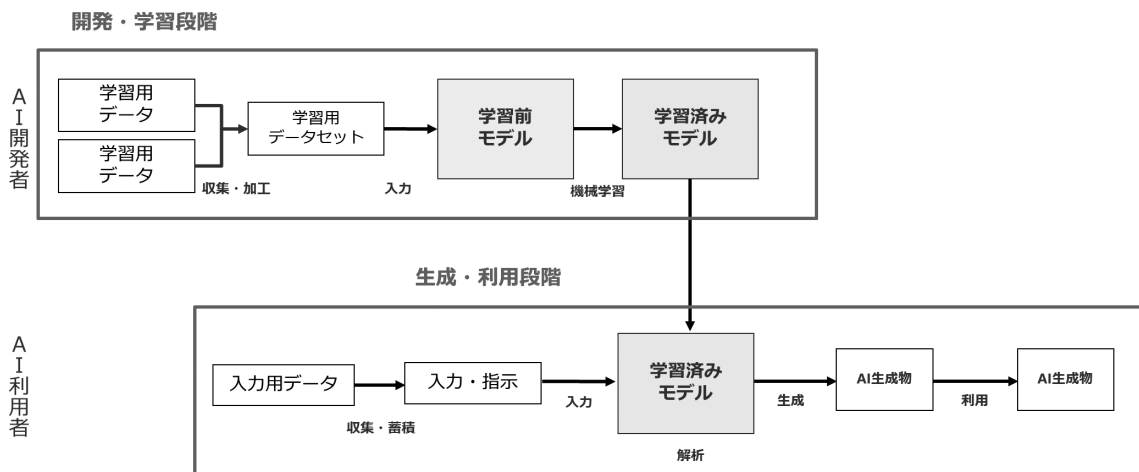


図 5

つまり、AI を開発するための行為が行われるフェーズを「開発・学習」段階、開発された AI を用いて実際に AI 生成物が生成されるフェーズを「生成・利用」段階と呼ぶこととする。

おそらく、一般的には「開発・学習」段階と「生成・利用」段階という用語は、このような意味で用いられていると思われる。このように整理すると、ここでいう「学習」とは、機械学習技術を用いて、AI のパラメータを物理的に更新する行為を指す。

一見当たり前のように見えるが、「開発・学習」段階と「生成・利用」段階の用語が、上記とは異なる意味で用いられることもあるため注意が必要である。

たとえば、「考え方」では、「開発・学習」段階における問題として、「生成 AI への指示・入力に用いるためのデータベースの作成に伴う著作物の利用行為」（「考え方」19 頁図 3）や、RAG（Retrieval-Augmented Generation）のことが紹介されているが（同 21 頁以下）、これらの問題は、図 5 の通常用語の意味に従うと「生成・利用」段階の問題である。

AI の「学習」には「機械学習技術を用いて、AI のパラメーターを物理的に更新する行為」という意味での「学習」に加えて、いわゆる In-Context Learning（ICL）の意味を含めて用いられることがある。ICL とは、LLM（Large Language Models、大規模言語モデル）へ入力するプロンプトとして、① 指示文（質問文）と共に、② 出力生成の際に参考にさせたいデータや、理想的な入出力例のデータを同時に追加入力することで、出力精度を上げる手法である（そのような追加データを全く入力しない手法を Zero-shot Learning、1 つだけ与える手法を One-shot Learning、いくつか与える手法を Few-shot Learning という⁽³⁾）。

「考え方」の整理を見ると、同「考え方」における「開発・学習段階」における「学習」とは、上記の「学習」（AI のパラメーター更新行為）だけではなく、そのようなパラメーター更新行為を伴わない ICL を含んだ概念なのではないかと思われる。

すなわち「考え方」における「開発・学習」段階と「生成・利用」の意味をまとめると以下の図 6 となる⁽⁴⁾。

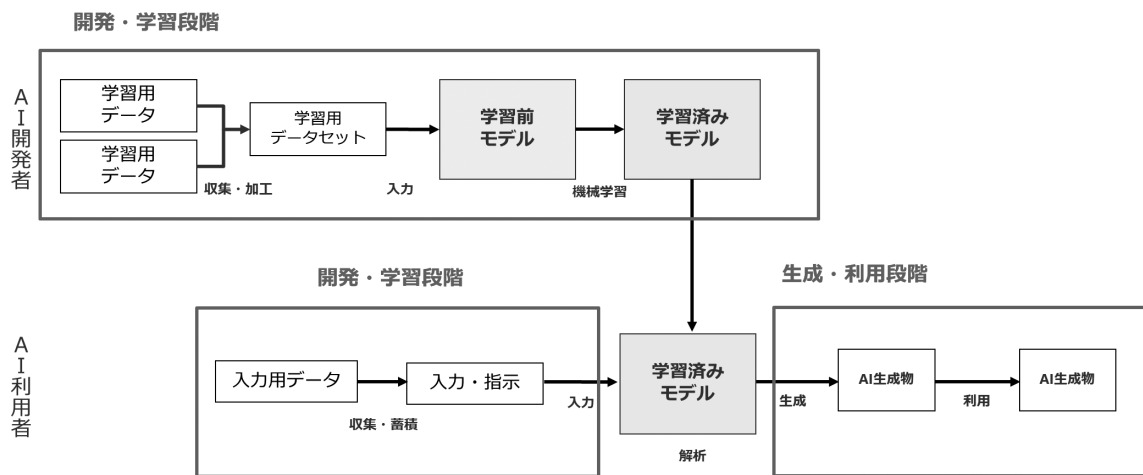


図 6

ただし、本稿では、これまで一般的に使われてきた用語との整合性という観点から、先ほど説明したような図 5 の意味で「開発・学習段階」と「生成・利用段階」の用語を用いる。

なお、「学習」も「In-Context Learning（ICL）」もいずれも「情報解析」（法第 30 条の 4 第 2 号）に該当する⁽⁵⁾ため、「開発・学習」段階、「生成・利用」段階をどのような意味で用いるかによって、法第 30 条の 4 や法第 47 条の 5 の適用関係についての結論が変わるわけではない。

- (3) たとえば LLM への入力プロンプトとして「・・・について説明して下さい。その際には以下の資料を参考して下さい。#資料1 #資料2」のように、指示文と共に、正確な出力生成に必要な文書等のデータを同時に入力する方法が ICL である。
- (4) このように考えると、「考え方」19 頁「図 3 生成 AI への指示・入力に用いるためのデータベースの作成に伴う著作物の利用行為」が「開発・学習段階における著作物の利用行為」として整理されていることも理解できる。「『AI と著作権に関する考え方について（素案）』に関するパブリックコメントの結果について」（以下、パブコメ）（https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_07/pdf/94011401_01.pdf）105、111、112、152 等も参照。
- (5) これは、「情報解析」が多種多様な情報処理を包含する広範な概念であること、また、ICL においては、モデル内のパラメーターの物理的な更新行為は行われませんが、見方を変えると、あたかもパラメーターを変えて学習した場合と同様に、指示や今生成しているデータにあわせて、パラメーターを一時的に更新し、モデルを急速に適応させているとみなすことができるとされていることを根拠とする。

1. 3 開発・学習段階と生成・利用段階を分けて検討する意味

AI と著作権については、開発・学習段階における著作物利用行為と生成・利用段階における著作物利用行為を分けて検討する必要があるが、これは著作権侵害の有無を検討する際には著作物の利用行為ごとに分けて検討するのが鉄則であることと、通常、開発・学習を行う行為者と生成・利用を行う行為者が別主体であるためである。

しかし、実は、両者を分けて検討する本当の意味は、著作権者にとっても、AI 開発・利用者にとっても、その影響範囲が格段に異なる点にある（表 1）。

表 1

	開発・学習段階	生成・利用段階
利用者（侵害者）による著作物の利用範囲（著作権侵害の範囲）	学習のための大規模な著作物利用（複製）行為	既存著作物と類似する個々の AI 生成物の生成・利用（複製）行為のみ
差止請求の範囲	大きい（一網打尽的・予防的）	小さい
損害賠償請求できる金額	大きい	小さい

特に、多数の著作物を、無断で学習に利用される可能性がある主体（報道機関等）にとっては、この点は大きなポイントと思われる。AI 開発者による学習自体が著作権侵害となると、学習に用いられた多数の記事単位で侵害責任を問えるからである。

一方、AI 開発者が、規範的行為主体として、生成・利用行為について行為主体（侵害主体）となることがあるが、そうなったとしても、あくまで生成された（複製された）著作物単位でしか侵害責任を問えない。

2. 生成 AI の開発・学習段階における著作物利用行為

2. 1 分析の視点

生成 AI の開発・学習段階における著作物利用行為について検討する際に、「考え方」においては、「【非享受目的】に該当する場合について」（19 頁以下）と【著作権者の利益を不当に害することとなる場合について】（22 頁以下）という 2 つの視点で分析をしている。

これは法第 30 条の 4 柱書但書の文言に忠実な視点だが、筆者は、生成 AI の開発・学習段階における著作物利用行為については、「学習目的による制限」と「学習対象による制限」に分けると、よりわかりやすいのではないかと考えている。

このうち、「学習目的による制限」が問題となるのは、主として、画像・音楽・動画等のコンテンツ生成 AI である。それらのコンテンツ生成 AI においては、特定の著作者の表現や作風・画風などを再現するニーズがあるためである。一方、LLM（Large Language Models、大規模言語モデル）のようなテキスト生成 AI においては、そのようなニーズは通常ないため「学習目的による制限」を考える必要はほとんどない。

一方、「学習対象による制限」については、コンテンツ生成 AI だけではなく、LLM のようなテキスト生成 AI の開発・学習にも影響が大きいといえる。

なお、ここでは「開発・学習段階における著作物利用行為のうち、どのような行為が著作権侵害に該当するのか」を検討しているが、実は「情報解析における著作物利用行為のうち、どのような行為が著作権侵害に該当するのか」という問いの方が問題の全体像を捉えている。

法第 30 条の 4 第 2 号や法第 47 条の 5 の条文では「開発・学習」ではなく「情報解析」という定義が用いられているため、正確には、「開発・学習」ではなく、「開発・学習段階における情報解析」と「生成・利用段階における情報解析」の双方を問題にする必要があるからである。

「開発・学習段階における情報解析と著作物利用行為」及び「生成・利用段階における情報解析と著作物利用行為」を図示すると以下のとおりとなる（図7及び図8）。

【開発・学習段階における情報解析】

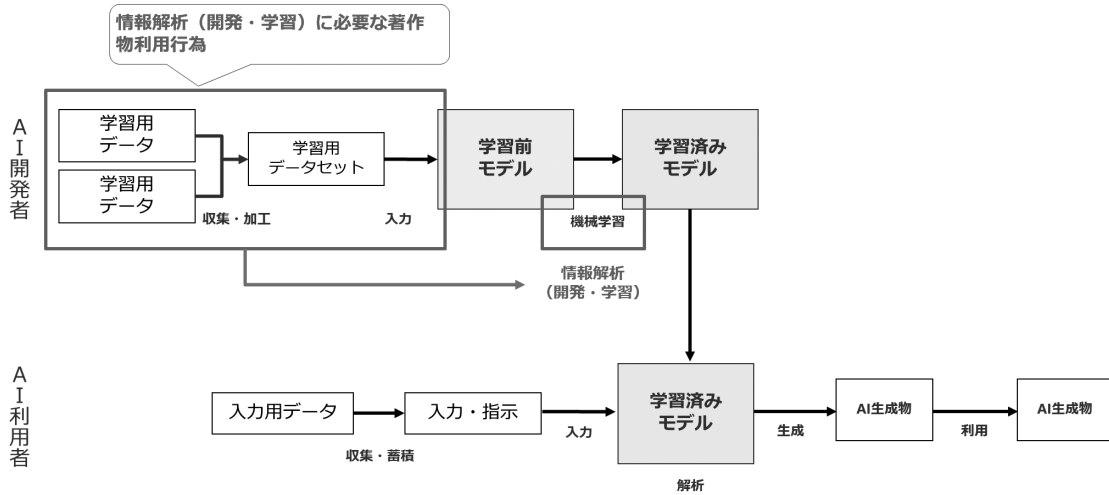


図7

【生成・利用段階における情報解析】

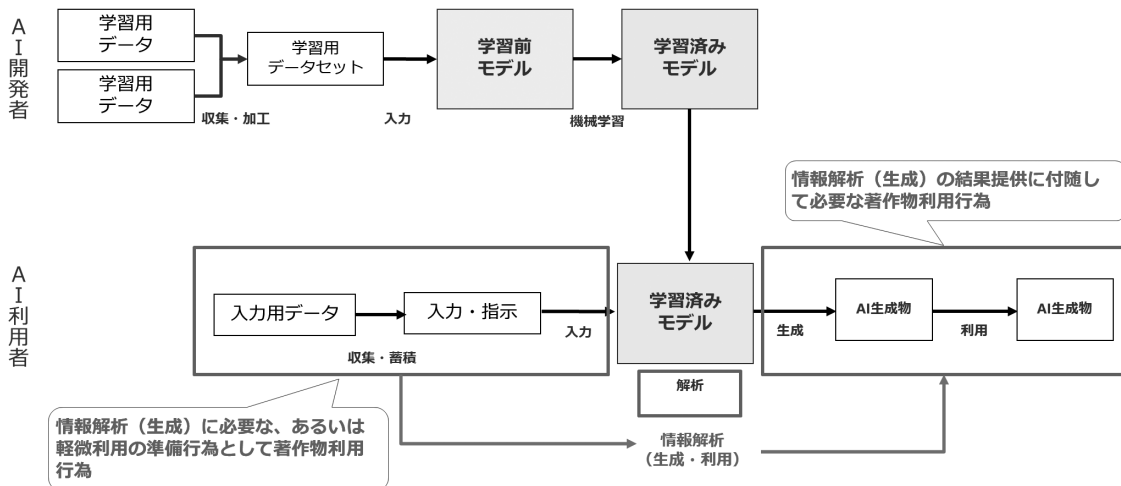


図8

その意味で、本来であれば「開発・学習段階及び生成・利用段階における情報解析に際しての著作物利用行為のうち、どのような行為が著作権侵害に該当するのか」を問題にし、「どのような目的で情報解析するか（情報解析目的）」と「何を情報解析するか（情報解析対象）」の方が正確な問題提起となるが、本稿では、まずは「開発・学習段階」にしばって検討し、「生成・利用段階」についてもそれが当てはまることについて後の「2. 5 生成・利用段階における情報解析と法第30条の4」で後述することとする⁽⁶⁾。

(6) 愛知靖之「日本法における権利制限～著作権法30条の4を中心に」上野達弘・奥邨弘司編著『AIと著作権』（勁草書房、2024年）13頁は「なお、30条の4などが適用され得る「情報解析」は、機械学習の場面（LoRAなどによる追加学習のケースも含む）のみならず、機械学習を経た学習済みモデルに著作物等のデータを入力し、コンテンツ等を生成させるという「推論」段階でも行われる。30条の4などの適用の可否は、学習段階・推論段階それぞれ個別に判断されることは言うまでもない。総じて、推論段階の方がAIによるコンテンツ生成により近い場面となるとは言えるが、30条の4など権利制限規定適用の可否については基本的に同様と考えられるため、以下では、原則両者を区別することなく、かつ学習段階を主な対象として検討を進める」とする。

2.2 学習目的による制限

(1) 表現出力目的

① 原則

法第 30 条の 4 柱書では、「次に掲げる場合その他の当該著作物に表現された思想又は感情を自ら享受し又は他人に享受させることを目的としない場合」と規定し、その上で、第 2 号において「情報解析（……）の用に供する場合」を挙げている。AI 学習行為は通常「情報解析」に該当することから、そのために行われる著作物利用行為は、原則として同条により適法となる。

② 表現出力目的とは

もっとも、対象著作物の利用に際して、「情報解析」のような非享受目的に加えて、当該著作物の享受目的が併存（享受目的併存）している場合には、法第 30 条の 4 は適用されない⁽⁷⁾。この点については、後述のように有力な反対説はあるものの、2018 年の著作権法改正当時の立法資料でも明らかにされており、現在の通説的な考え方と思われる⁽⁸⁾。

この「享受目的併存」する著作物利用行為には様々なタイプがある⁽⁹⁾が、生成 AI の学習においては、たとえば以下のような場合が想定される（「考え方」20 頁）。

既存の学習済みモデルに対する追加的な学習（そのために行う学習データの収集・加工を含む）のうち、意図的に、学習データに含まれる著作物の創作的表現の全部又は一部を出力させることを目的とした追加的な学習を行うため、著作物の複製等を行う場合。

（例）AI 開発事業者又は AI サービス提供事業者が、AI 学習に際して、いわゆる「過学習」（overfitting）を意図的に行う場合

したがって、このような「表現出力目的」での学習⁽¹⁰⁾には法第 30 条の 4 は適用されず、他の権利制限規定が適用されない限り著作権侵害に該当する。

なお、「情報解析」目的に加えて享受目的が併存している場合に法第 30 条の 4 が適用されないという上記通説には、条文の規定ぶり（法第 30 条の 4 柱書の「その他の」）等を根拠とした有力な反対説がある⁽¹¹⁾。

この反対説によると、「情報解析」に該当すれば、享受目的が併存していても法第 30 条の 4 本文が適用されることになるが、同説は、享受目的（表現出力目的）が併存する学習には、法第 30 条の 4 柱書但書に該

(7) 奥邨弘司「生成 AI と著作権に関する米国の動き－AI 生成表現の著作物性に関する著作権局の考え方と生成 AI に関する訴訟の概要」コピライト 63 巻 747 号（2023 年）46 頁、前田健「柔軟な権利制限規定の設計思想と著作権者の利益の意義」田村善之編著『知財とパブリック・ドメイン第 2 巻著作権法篇』（勁草書房、2023 年）206 頁

(8) 文化庁著作権課「デジタル化・ネットワーク化の進展に対応した柔軟な権利制限規定に関する基本的な考え方（著作権法第 30 条の 4、第 47 条の 4 及び第 47 条の 5 関係）」8 頁

(9) 享受目的が併存している著作物利用行為には、「享受」の時点に応じて 2 種類ある。1 つは対象著作物の利用行為と「同時に」対象著作物の享受が行われるパターンである。文化庁著作権課「デジタル化・ネットワーク化の進展に対応した柔軟な権利制限規定に関する基本的な考え方」8 頁で挙げられている例（家電量販店等においてテレビの画質の差を比較できるよう市販のブルーレイディスクの映像を常時流す行為（上映）、漫画の作画技術を身につけさせることを目的として、民間のカルチャー教室等で手本とすべき著名な漫画を複製して受講者に参考とさせるために配布したり、購入した漫画を手本にして受講者が模写したり、模写した作品をスクリーンに映してその出来映えを吟味してみたりするといった行為）はいずれもこのパターンである。もう一つは、対象著作物の利用行為の「後」で対象著作物の享受が行われるパターンである。「生成 AI における表現出力目的の学習」の問題はこのパターンである。生成 AI の学習段階における著作物の「情報解析」のための利用行為の際に、その「後」の生成・利用段階で生じる享受行為の目的があるかを問題とするためである。

(10) もちろん、表現出力目的での「生成」にも法第 30 条の 4 は適用されない（「考え方」22 頁、37 頁～38 頁）。表現出力目的での「情報解析」に法第 30 条の 4 が適用されないためである。

(11) 前掲注 6・15 頁

当とする⁽¹²⁾ため、結論にはあまり相違はない。

③ どのような場合に「表現出力目的」が認められるのか

「考え方」は、「表現出力目的」がある具体例として、先述のように「(例) AI 開発事業者又は AI サービス提供事業者が、AI 学習に際して、いわゆる「過学習」(overfitting) を意図的に行う場合」を挙げる。

これはたとえば、以下のようなケースを指している (図 9、図 10)^{(13)、(14)}。

【AI 開発事業者の学習における表現出力目的】

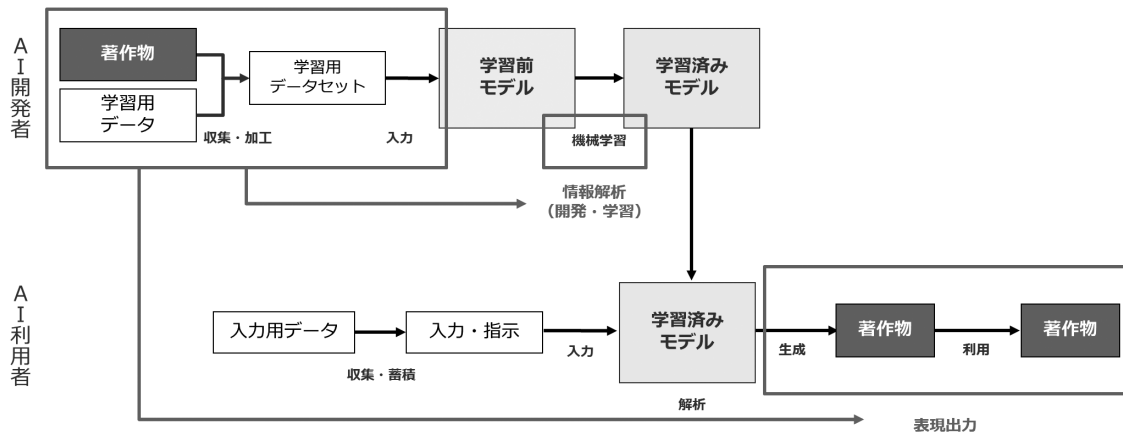


図 9

【AI サービス提供者の追加学習における表現出力目的】

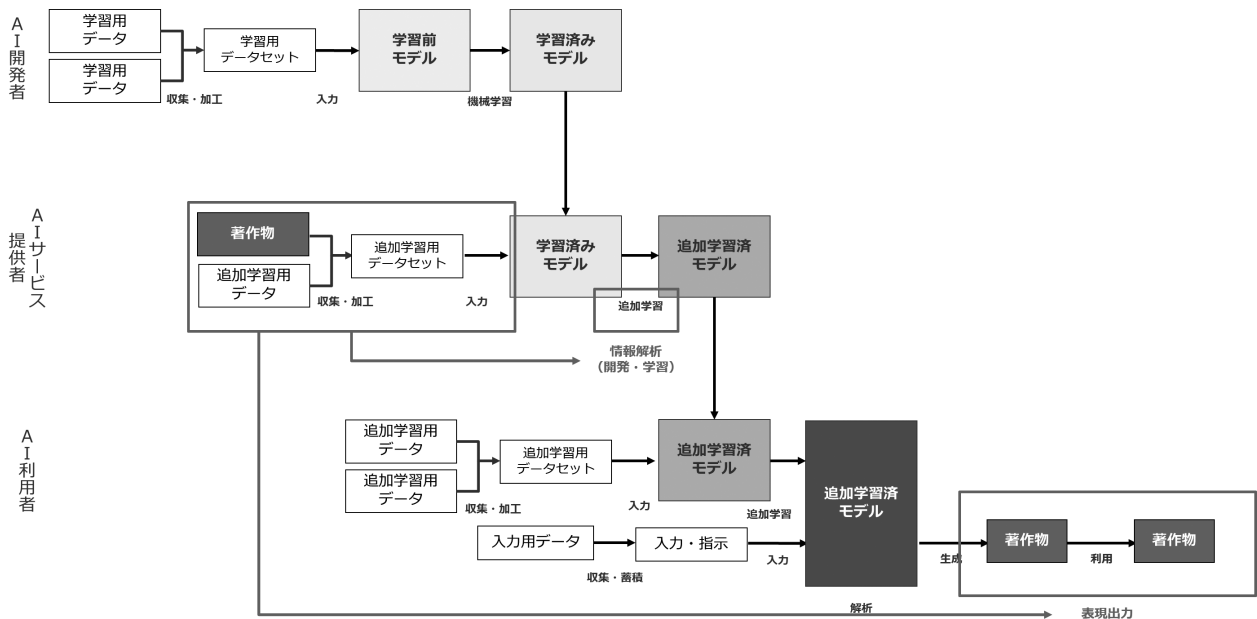


図 10

- (12) 前掲注 6・38 頁、208 頁以下。ただし、正確には愛知教授は、表現出力目的がある場合のみならず、本文で後述するように、「作風」を再現する目的がある場合も法第 30 条の 4 柱書但書に該当するとする。
- (13) なお、享受目的には、自らが享受する目的だけでなく、第三者に享受させる目的も当然含む (法第 30 条の 4 柱書は「次に掲げる場合その他の当該著作物に表現された思想又は感情を自ら享受し又は他人に享受させることを目的としない場合」とする)。
- (14) 表現出力「目的」の情報解析が法第 30 条の 4 の適用外なので、理論的には学習対象となった既存著作物が実際に出力されなくても当該「目的」が認定される可能性もあるが、実際には後述のように当該出力が頻発しないと「目的」が認定される可能性はないと思われる。

図 9 は、AI 開発者が、AI 開発という情報解析目的に加えて、「AI 利用者において学習対象データと同一・類似の AI 生成物を出力させる目的」で、AI 学習に際して過学習を意図的に行うケースである。

図 10 は、AI サービス提供者が、AI 開発という情報解析目的に加えて、「AI 利用者において学習対象データと同一・類似の AI 生成物を出力させる目的」で、AI の追加学習に際して過学習を意図的に行うものである。

また、AI 利用者自身の追加学習において表現出力目的があるケースも当然想定される。たとえば、以下のように AI 利用者自身が表現出力目的で過学習を意図的に行う場合である（図 11）。

【AI 利用者の追加学習における表現出力目的】

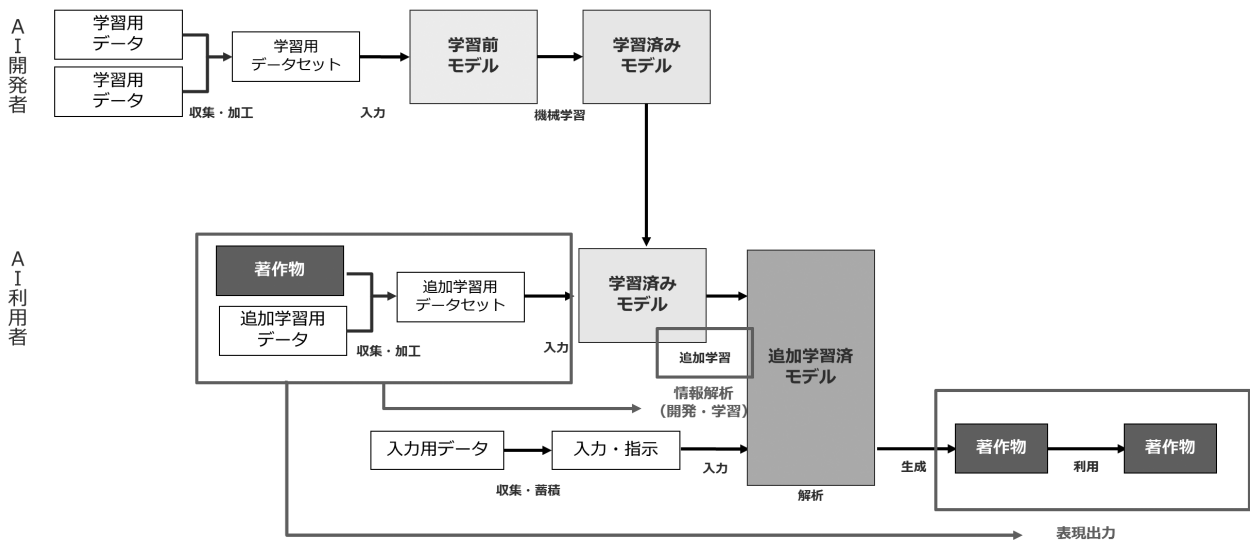


図 11

もっとも、この「目的」をあまりに広く捉えると、情報解析が過度に制限され法第 30 条の 4 の趣旨が没却されかねない。そのため、汎用的な生成 AI において、学習用データと類似性の認められる出力物が結果として出力される可能性を認識・認容していたという程度であれば、享受目的は認められないと考える⁽¹⁵⁾。

④ 「表現出力目的」の立証方法

「表現出力目的」がある場合には法第 30 条の 4 は適用されないが、その場合であっても後述のように法第 47 条の 5 は適用可能である。もっとも、後述のように法第 47 条の 5 が適用されるには「付随性」「軽微利用」などの非常に厳しい要件を満たす必要があり、現実には法第 47 条の 5 が適用される場面はかなり限られていると思われる。

したがって、結局のところ「表現出力目的」の有無が、学習段階における著作物の利用行為の適法性判断のための重要なメルクマールとなる⁽¹⁶⁾。

もっとも「表現出力目的」というのは主観的な要素であることと、「開発・利用段階」より「後」の段階である「生成・利用段階」における目的の有無（つまり、まだ生じていない表現の出力行為を行う目的の有無）を問題にするため、一般的にはその立証が難しいといえる。

そのため、実際には、「表現出力目的」の存在を推認させるような客観的な事実（間接事実）があるかを検討することになる（「考え方」20 頁脚注 23）。

(15) 前田健「生成 AI の利用が著作権侵害となる場合」法学教室 No.523（2024 年）29 頁

(16) 前掲注 15・29 頁は「享受目的の有無が、学習用データとしての利用が認められるかどうかを分かつ、重要なメルクマールとなる。」とする。

まず、生成・利用段階において、学習対象著作物と創作的表現が共通した生成物が生成される事例があるという事実だけでは、開発・学習段階における享受目的の存在を推認することまではできない（「考え方」21 頁）。生成 AI の技術的な特質から、そのような事態を完全に防止することは困難であり、結果としてそのような事態が生じたとしても、「表現出力目的」がないこともあるからである。

一方、生成・利用段階で「学習された著作物と創作的表現が共通した生成物の生成が著しく頻発するといった事情」があれば、「表現出力目的」が推認されよう（「考え方」21 頁）⁽¹⁷⁾。

さらに、事業者が、学習対象著作物の表現がそのまま出力されやすいような特殊な学習手法（特定の作品や特定の著作者の作品のみを集中的に学習させていること）を採用している場合、表現出力目的を推認させる間接事実となる。

一方、侵害物の生成を抑止するための実効的な技術的手段を講じている場合、当該事業者の行う AI 学習のための複製が、非享受目的である（＝表現出力目的がない）ことを推認させる事情となり得る（「考え方」20 頁脚注 22）⁽¹⁸⁾。表現出力目的があるのであれば、そのような技術的手段を講じるはずがないからである。

このような様々な要素を総合的に検討しながら生成 AI の開発・学習段階における著作物利用行為の「表現出力目的」の有無が判断されることとなる⁽¹⁹⁾。

（2）作風模倣目的

次に、特定の著作権者の作風模倣目的での開発・学習における著作物利用行為、すなわち「特定のクリエイターの作品である少量の著作物のみを学習データとして追加的な学習を行うことで、当該クリエイターの「作風」を模倣する目的で学習を行うこと」について検討する。

（ア）学説の状況

① 法第 30 条の 4 柱書但書該当説

まず、作風模倣目的での開発・学習における著作物利用行為は、法第 30 条の 4 柱書但書に該当するという説がある⁽²⁰⁾。具体的には、「ディズニー映画風の新しい映画を作る AI を開発するために、ディズニー映画（但書にいう「当該著作物」）全てをコンピュータに入力して機械学習させる行為」のように、特定著作権者の著作物を狙い撃ちし、当該著作権者の「作風」を備えることで市場競合するコンテンツを生成するために、その著作物の全てあるいはこれを大量に学習・推論に利用する行為が法第 30 条の 4 柱書但書に該当するとする。

この説は、法第 30 条の 4 が、情報解析のみならず、非享受利用一般を対象とする包括的な規定となったことに伴って、但書も包括的な文言となったこと、特定の作者固有のオリジナリティが認められ、その者の作品であると容易に同定し得るような具体性及び表現との近接性・連続性を備えた「作風」を模倣する目的で当該作品を大量に学習する行為は、当該作品に関する「将来における潜在的販路」を害し「潜在的な市場競合」をもたらすことを根拠として、当該行為が法第 30 条の 4 柱書但書に該当するとする。

② 法第 30 条の 4 柱書但書非該当説

これに対して、法第 30 条の 4 柱書但書の適用対象は限定されるとの立場を前提とした上で、但書該当説

(17) ただし、当該頻発が、生成 AI の利用者が既存の著作物の類似物の生成を意図して生成 AI に入力・指示を与えたこと等に起因するものである場合は、このような事情があったとしても、AI 学習を行った事業者の享受目的の存在を推認させる要素とはならない（「考え方」21 頁脚注 24）。

(18) 前掲注 4・パブコメ 165

(19) 前田教授は非享受目的について、「要件事実に言えば、規範的要件のようなところがある」とする（前掲注 7・233 頁）。

(20) 前掲注 6・32 頁、金子敏哉「生成 AI とせいせいしない著作権法」法学セミナー 828 号 51 頁（2024 年）58 頁

を批判する見解がある。

まず筆者は、但書該当説がその根拠とする「将来における潜在的な販路を害する」かどうかは、未だ実際の AI 生成物が生成されていない段階での判断であることからすると、かなり曖昧であること、そもそも「対象著作物の利用行為とは異なる、将来発生する可能性のある著作物の利用行為」についても対象著作物の利用行為の但書該当行為の判断に織り込むというのは、但書該当行為を広く解しすぎ、結果として潜在的な市場を害さず権利制限行為に該当する（著作権侵害に該当しない）行為を、著作権侵害行為とするものであり、法第 30 条の 4 の趣旨を没却しかねないと批判した⁽²¹⁾。

また、但書で保護される「著作権者の利益」は、あくまで著作権によって保護される利益のはずであって「作風」や「画風」のように著作権で保護されないものが共通することで生じる市場競合によって害される利益は含まれないとの批判もある⁽²²⁾。

（イ）検討

この点に関しては、まず「作風」は通常は「アイデア」に属するものであることからすると、生成・利用段階において、仮に既存著作物と「作風」が類似する AI 生成物を生成・利用したとしても、当該既存著作物の著作権侵害には該当しない。

但書該当説は、学習対象著作物と作風が類似する AI 生成物を生成する目的での著作物利用行為を著作権侵害とするものであるが、学習対象著作物と作風が類似する AI 生成物が生成・利用されても著作権侵害にならないにもかかわらず、なぜその前の段階（「作風類似の AI 生成物すら生成されていない段階」）における著作物利用行為が著作権侵害になるのかについての明確な説明がつかないように思われる。

また、但書該当説は、学習対象著作物と作風が共通する AI 生成物が生成・利用されることにより学習対象著作物の「将来における潜在的な販路を害する」とするが、そのような解釈は、実質的には作風（アイデア）を著作権法により保護することを意味することになる。

さらに、但書該当説は、同説の実質的な根拠として、「(多数の著作権侵害) コンテンツ生成・利用の未然防止を図る」ことを強調する⁽²³⁾。

そのような必要性があることは否定できないが、著作権侵害コンテンツの生成・利用という著作物利用行為と、それに向けられた開発・学習における著作物利用行為は全く別の著作物利用行為である。

「(多数の著作権侵害) コンテンツ生成・利用の未然防止を図る」ためには、**既存著作物と同一・類似の AI 生成物の生成・利用行為により著作権侵害が実際に生じた場合において**、当該侵害行為に対する権利者の差止請求としてどの範囲の差止請求を認めるべきか、という観点から検討すべき問題であって、開発・学習における著作物利用行為を著作権侵害とすべきということには結びつかないと考える。

したがって、作風模倣目的での学習対象著作物の利用は法第 30 条の 4 柱書但書には該当しないと考える。

もっとも、実際には、模倣の対象となっている「作風」が「アイデア」に該当するのか「表現」に該当するのかは、非常に区別が難しい（「考え方」21 頁、同 24 頁）。

したがって、「作風模倣目的」が「表現出力目的」に該当して法第 30 条の 4 の適用外となるのか、あるいは「アイデア」の模倣目的に該当して法第 30 条の 4 により適法となるかは、ケースバイケースであろう。

(21) 柿沼太一「生成 AI と著作権」知財・ふりずむ 21 巻 248 号（2023 年）19 頁～20 頁、奥邨前掲注 7・46 頁～47 頁

(22) 奥邨前掲注 7・47 頁

(23) 前掲注 6・31 頁において愛知教授は「いったんディズニー映画風の新しい映画を作る AI が開発され、ユーザーに提供されると。それを利用したコンテンツ生成及びその後の利用が多数集積し、著作権者に対して大きな経済的不利益が生じ得る」とする。

(3) まとめ

以上のように、開発学習段階における著作物利用行為のうち「学習目的による制限」としては、「表現出力目的」がある著作物利用行為のみが（他の権利制限規定の適用がなければ）著作権侵害に該当し、「作風模倣目的」があるにすぎない著作物利用行為は著作権侵害に該当しない。

「表現出力目的」や「作風模倣目的」が、実際に問題となるのは、画像・動画・音楽・音声生成 AI のようなコンテンツ生成 AI に限定されるであろう。そのようなコンテンツ生成 AI の学習においては、特定のクリエイターの表現・作風再現を目的とすることがあるからである。

一方、LLM のようなテキスト生成 AI の学習においては、そのような「表現出力目的」や「作風模倣目的」を有することは通常はないため、学習目的による制限が問題となる場面はほとんどないと思われる。

2. 3 学習対象による制限

(1) 情報解析に活用できる形で整理したデータベースの著作物

① 情報解析用 DB 著作物とは

法第 30 条の 4 柱書但書に該当する例として、同条制定当時から、「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物が販売されている場合に、当該データベースを情報解析目的で複製等する行為」が示されている⁽²⁴⁾（「考え方」24 頁。この「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物」を、以下「情報解析用 DB 著作物」という）。

情報解析用 DB 著作物が販売されている場合において、それを学習（情報解析）目的で利用したり、生成・利用段階で、AI による解析（情報解析）目的で AI に入力したりする行為は、法第 30 条の 4 柱書但書に該当し、他の権利制限規定が適用されなければ著作権侵害に該当する。

もっとも、注意しなければならないのは、「考え方」24 頁～27 頁で「(ウ) 情報解析に活用できる形で整理したデータベースの著作物の例について」「(エ) 本ただし書に該当し得る上記 (ウ) の具体例について（学習のための複製等を防止する技術的な措置が施されている場合等の考え方）」に記載されているのは、この「情報解析用 DB 著作物」に関するものであり、それ以外の著作物に関するものではない、ということである。

上記「(ウ)」「(エ)」においては、報道機関がウェブ上で個々の記事を提供しているケースについて記載されているが、これは当該個々の記事の著作権侵害のことを問題にしているわけではない。

つまり「考え方」の当該部分は、「個々の記事をクロールした結果、報道機関が別途提供している（あるいは将来提供予定の）情報解析用 DB 著作物の複製に該当する場合には、当該情報解析用 DB 著作物の著作権侵害になる」と述べているだけであって、あくまで、当該情報解析用 DB 著作物の著作権侵害を問題にしているにすぎない。

これは「考え方」26 頁の「○ そのため、AI 学習のための著作物の複製等を防止する技術的な措置が講じられて……」の部分において、「この措置（筆者注：複製防止のための技術的措置）を回避して、クローラにより当該ウェブサイト内に掲載されている多数のデータを収集することにより、AI 学習のために当該データベースの著作物の複製等をする行為」とされていることから明らかである。

また、小委員会での議論で上野委員ら複数の委員からこの点について指摘があり、事務局から明確に以下のような回答がなされている（第 6 回小委員会議事録・上野委員、澤田委員・中川委員発言参照）。

(24) ただし、「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物が販売されている場合に、当該データベースを情報解析目的で複製等する行為」については、法第 30 条の 4 柱書但書の問題ではなく、当該データベース著作物の本来的利用であるとして享受目的併存により法第 30 条の 4 が適用されない、という説もある（前掲注 15・30 頁、前掲注 7・232 頁）。

【三輪著作権課調査官】事務局でございます。ただいま委員の先生方からいただきました意見を踏まえまして少し補足させていただきますと、今おっしゃっていただきましたように、この点、記載の趣旨としては、先ほど事務局から御説明申し上げたとおり、対象の著作物として考えておりますのは、情報解析用のデータベースの著作物であり、問題にする行為としても、情報解析用のデータベースの著作物の複製と言えるような行為については、30条の4ただし書に該当し、権利制限の対象とはならない場合があると、そういう趣旨の記載をしているというところでございます。

② 情報解析用 DB 著作物の複製が行われる具体例

「考え方」24頁～27頁で示されている、情報解析用 DB 著作物の複製が行われる具体例は以下の3つである。

▼ 具体例 1

情報解析用 DB 著作物が DVD 等の記録媒体に記録して提供されている場合にそのまま当該記録媒体から複製する行為

▼ 具体例 2

インターネット上のウェブサイトで、ユーザーの閲覧に供するため記事等が提供されているのに加え、情報解析用 DB 著作物が API を通じて有償で提供されている場合において⁽²⁵⁾、当該 API を有償で利用することなく、当該ウェブサイト閲覧用に掲載された記事等のデータから、当該データベースの著作物の創作的表現が認められる一定の情報のまとまりを情報解析目的で複製する行為

▼ 具体例 3

AI 学習のための著作物の複製等を防止する技術的な措置が講じられており、かつ、このような措置が講じられていることや、過去の実績（情報解析に活用できる形で整理したデータベースの著作物の作成実績や、そのライセンス取引に関する実績等）といった事実から、当該ウェブサイト内のデータを含み、情報解析に活用できる形で整理したデータベースの著作物が将来販売される予定があることが推認される場合に、この措置を回避して、クローラにより当該ウェブサイト内に掲載されている多数のデータを収集することにより、AI 学習のために当該データベースの著作物の複製等をする行為

(i) 具体例 1

このうち具体例 1 については、改正前法第 47 条の 7 の但書に該当する行為であって、情報解析用 DB 著作物の複製として法第 30 条の 4 柱書但書に該当することは明らかである（図 12）。

(25) なお「考え方」25頁脚注 28 には「この点に関して、インターネット上のウェブサイトに掲載されたデータについては、AI 学習のための複製を行うクローラによるウェブサイト内へのアクセスが、後述するウェブサイト内のファイル”robots.txt”への記述により制限されていない場合、「(大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物が) 販売されている場合」に該当しないことを推認させる要素となるものと考えられる」との記載がある。この記載内容は合理的であるが、このような技術的制限が講じられていない事実は、パターン 3 における「情報解析に活用できる形で整理したデータベースの著作物が将来販売される予定があること」に該当しないことを推認させる要素にもなると思われる（前掲注 4・パブコメ 191 参照）。

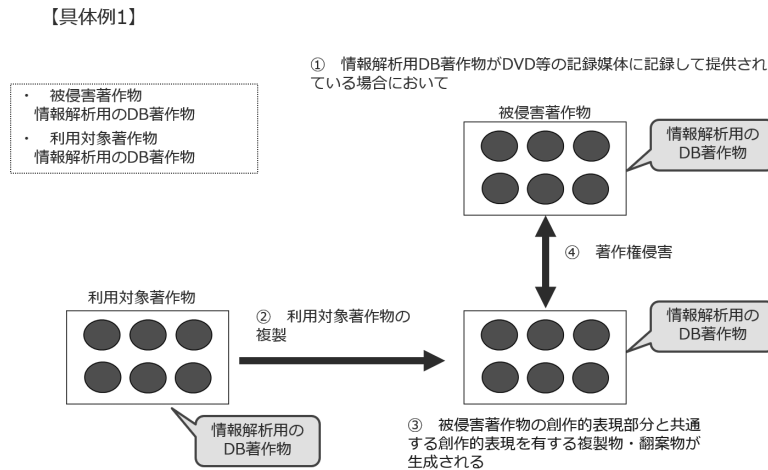


図 12

一方、具体例 2 及び具体例 3 は、以下に述べるように、一言で言うと「**理論的にはあり得るが、発生する可能性がほぼないケース**」ではないかと思われる。

(ii) 具体例 2

具体例 2 で権利者が著作権侵害を主張する場合、権利者（報道機関）が主張立証しなければならないのは以下の事実である。

▼ 自社が情報解析用 DB 著作物を API を通じて提供していること

まず、単なる DB を提供しているだけではこの要件を満たさない。後述するように、情報解析用 DB 著作物に該当しない、単なる DB についてのライセンス市場がある場合であっても、当該 DB を情報解析の対象とすることは法第 30 条の 4 柱書但書に該当しないからである。

また、仮に当該 DB が DB 著作物だとしても、当該 DB 著作物は「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物」（情報解析用 DB 著作物）でなければならない。

たとえば、印刷用やウェブサイトを利用するためのイラストや画像で構成される DB 著作物については、それを情報解析に利用したとしても、当該 DB 著作物の市場を害することはないため、法第 30 条の 4 柱書但書には該当しない⁽²⁶⁾。要するに DB 著作物であっても、「情報解析に使おうと思えば使える」だけでは足りず、「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物」である必要がある⁽²⁷⁾。

報道機関の提供している記事 DB は、通常は「ある一定の期間の記事を網羅的に収録した DB」であると思われるが、その場合、そもそも「情報の選択」についての創作性がないことから、DB 著作物、ましてや情報解析用 DB 著作物には該当しない可能性が高い。

▼ AI 開発者が個別の記事を集積した結果、権利者が販売している情報解析用 DB 著作物の創作的表現部分を利用するに至ったこと

また、具体例 2 で著作権侵害が認められるには「AI 開発者が個別の記事を集積した結果、権利者が API を通じて有償で提供している情報解析用 DB 著作物の創作的表現部分を利用するに至ったこと」が必要である（図 13）。

(26) 前掲注 6・座談会奥郵発言（231 頁）は「例えば、一般的な印刷用とかウェブサイトの挿絵用とかでライセンスされているフォトストックのデータベースだと著作物の種類はデータベースでも、用途は観賞、利用態様は、情報解析用機器で全部複製するという当てはめになります。これはもともと、柱書本文がやってもいいよと書いてあることをやっているだけなので、不当に害する余地はない、大丈夫じゃないかなと私は思います。」とする。

(27) 前田健「生成 AI における学習用データとしての利用と著作権」（有斐閣オンライン）脚注 16

【具体例2】

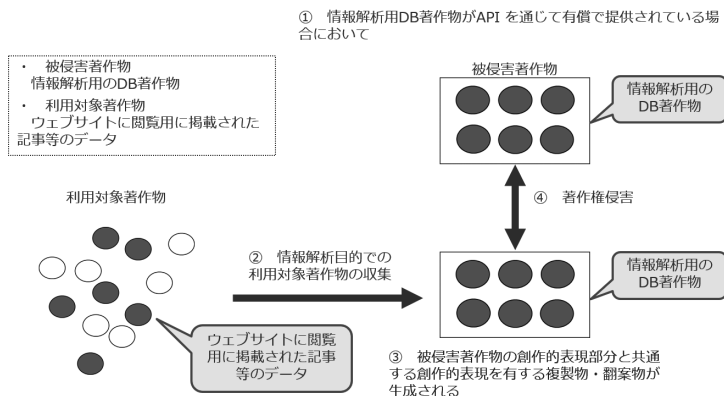


図 13

つまり、AI 開発者が個別の記事を集積したとしても、それが、権利者が API を通じて有償で提供している情報解析用 DB 著作物の創作的表現部分を利用していなければ、当該情報解析用 DB 著作物の著作権侵害にはならない。

そして、報道機関の提供している記事 DB が情報解析用 DB 著作物に該当する可能性があるとする、「世の中のありとあらゆる事実関係の中から、報道に値する事実関係を選択した」という点に「情報の選択」における創作性が認められる場合ではないかと思われる⁽²⁸⁾。

もっとも、AI 開発者は、AI 開発のために複数の報道機関のニュース記事や他のウェブ上の文章をまとめて大規模に収集するのが通常であるから、1つの報道機関だけの、かつ当該報道機関のみが報道しているニュースのみをあえて対象にして収集して学習に用いることは考えにくい。そのように、AI 開発者が、複数の報道機関のニュース記事や他のウェブ上の文章をまとめて大規模に収集して学習に用いた場合、ある報道機関の情報解析用 DB 著作物において、記事の選択に創作性が認められるとしても、当該創作的表現は利用されていないことになる（全体比較論）。

▼ まとめ

つまり、具体例2において著作権侵害に該当するのは、ウェブサイトに閲覧用に掲載された記事等のデータを API を経由せずに収集する AI 開発者が、あえて「情報解析用 DB 著作物」と同じ「情報の選択の創作性」を有する DB の著作物を作成するという極めて「偶然」の場合のみである。

このようなことは現実的にはほぼあり得ないであろう。

そもそも、AI 開発者としては、そのような行為をあえて行う必要性が全くないからである。

(iii) 具体例3

具体例3（図14）の場合「将来販売される可能性がある情報解析用 DB 著作物」の将来販路が阻害されることが必要だが、そのような事態は、具体例2以上に「現実的にはあり得ない」と言って良いと考える。

(28) ただし、そう考えるとしても、他の新聞社も同様の事実関係を報道していればそのような「情報の選択」についての創作性はない。

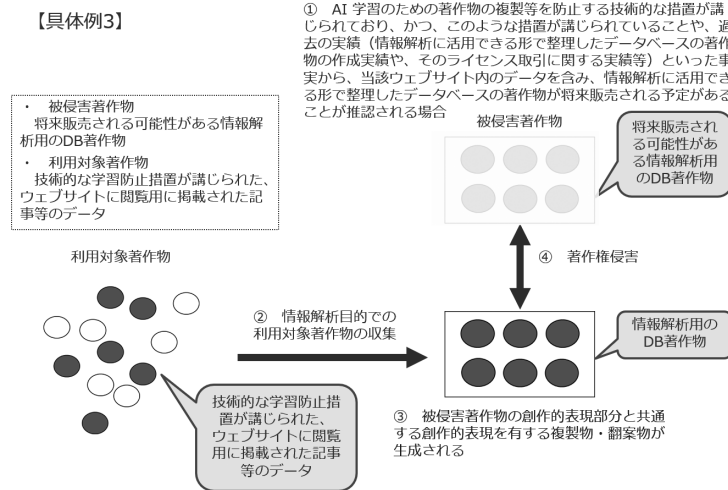


図 14

なぜなら、具体例 2 の場合には、「現実に販売されている情報解析用 DB 著作物」がある場合において、当該情報解析用 DB 著作物の創作的表現部分を利用する行為を対象にしていたが、具体例 3 の場合には「将来販売される可能性がある（言い換えれば、まだ世の中に存在していない）情報解析用 DB 著作物」を対象とするものだからである。

このように「将来販売される可能性がある情報解析用 DB 著作物」の将来販路が阻害されることを理由に個々のウェブ記事の収集行為が法第 30 条の 4 柱書但書に該当するという「考え方」の解釈は、実質的には、現時点では存在しない、将来発生する可能性のある著作物についての著作権侵害を認める解釈である。

確かに、著作物自体が未だに発生しておらず、かつ侵害行為も行われていない段階での予防請求を認めた裁判例は存在する（東京地判平 5・8・30 知的裁集 25 卷 2 号 380 頁（ウォール・ストリート・ジャーナル事件））。

この判決は、米国において日刊新聞「The Wall Street Journal」（「本件新聞」）を継続して発行する X（債権者・被控訴人）が、わが国において本件新聞の記事を抄訳して紙面構成に対応して配列した文書（以下「本件文書」という）を募集した会員に作成・頒布する Y（債務者・控訴人）に対し、本件文書の作成・頒布は、債権者の本件新聞について有する編集著作権を侵害するとして申し立てた本件文書の作成・頒布の差止め処分が認められた事案である。

この事件では、将来作成される著作物の編集著作権に基づく差止めの可否等が問題となり、差止めが認められた。

しかし、この事件は、過去に具体的な編集著作権侵害行為が継続して行われていたことを根拠として、将来的に同様の著作権侵害が発生する可能性が相当高度であることを認め、その結果差止請求を認めたものにすぎず、一般化することはできない。現に同裁判例については、「被侵害著作物が未だ存在しない場合の差止請求は極めて例外的であり、同判決は一般化できないであろう」という評価がされている⁽²⁹⁾。

したがって、同裁判例を根拠として「将来発生する可能性のある著作物についての潜在的販路の阻害」行為が法第 30 条の 4 但書に該当するとすることはできないと考える。

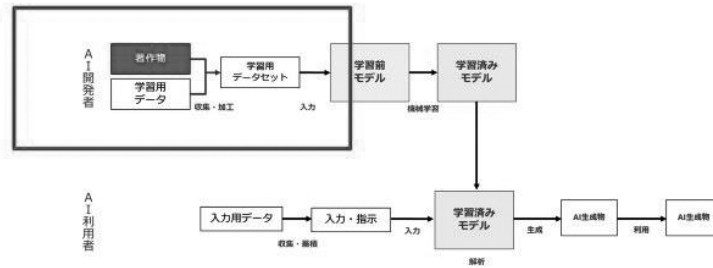
(2) 海賊版等の権利侵害複製物

AI 開発者が、海賊版等の権利侵害複製物を開発・学習段階で収集・利用した場合、著作権侵害責任を問

(29) 中山信弘『著作権法 第 3 版』（有斐閣、2020 年）727 頁

われることがあり得るが、それが「開発・学習行為」に関する責任なのか「生成・利用行為」に関する責任なのかを明確に区別する必要がある（図 15）。

① AI開発者が海賊版等をWEBサイトから収集して開発・学習に利用する行為



② AI開発者が海賊版等を学習対象として生成AIを作成・提供し、同生成AIを利用したAI利用者が当該海賊版と同一・類似の生成物を生成・利用する行為

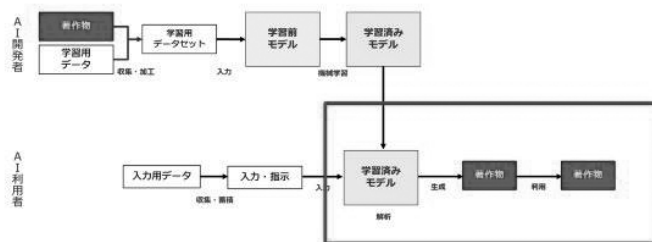


図 15

① AI開発者が海賊版等をウェブサイトから収集して開発・学習に利用する行為

諸外国の情報解析規定においては、情報解析の客体が適法にアクセスしたものであることを条件とした立法例が多く（英国法、欧州 DSM 指令、スイス法、シンガポール法）、そのような立法の下では、AI 開発者が海賊版等をウェブサイトから収集して開発・学習に利用した場合、当該開発・学習行為自体が著作権侵害になる。

一方、日本国著作権法の法第 30 条の 4 にはそのような例外が設けられていない。

さらに、法第 30 条第 1 項第 3 号や、法第 47 条の 5 第 1 項但書においては、権利侵害複製物の利用行為について権利制限の対象外であることが明記されているが、法第 30 条の 4 にはそのような規定が存在しないことと対比して考えると、AI 開発者が海賊版等をウェブサイトから収集して開発・学習に利用する行為についても、法第 30 条の 4 が適用されて原則として適法となると考える⁽³⁰⁾。

② AI開発者が海賊版等を学習対象として生成 AI を作成・提供し、同生成 AI を利用した AI 利用者が当該海賊版と同一・類似の生成物を生成・利用する行為

一方、①とは別の、生成・利用段階の問題として、AI 開発者が海賊版等を学習対象として生成 AI を作成・提供し、同生成 AI を利用した AI 利用者が当該海賊版と同一・類似の生成物を生成・利用した場合に、AI 開発者が同生成・利用についてどのような責任を負うか、という問題がある。

この点については、一定の場合には、当該生成・利用に用いられた AI の開発者・サービス提供者が同生成・利用による著作権侵害行為の規範的な主体として責任を負う場合があり得る。

「考え方」28 頁の以下の記載はその点を説明したものである。

(30) 前掲注 6・66 頁

○ 特に、ウェブサイトが海賊版等の権利侵害複製物を掲載していることを知りながら、当該ウェブサイトから学習データの収集を行うといった行為は、厳にこれを慎むべきものである。この点に関して、生成・利用段階においては、後掲(2)キのとおり、既存の著作物の著作権侵害が生じた場合、AI 開発事業者又は AI サービス提供事業者も、当該侵害行為の規範的な主体として責任を負う場合があり得る。この規範的な行為主体の認定に当たっては、当該行為に関する諸般の事情が総合的に考慮されるものと考えられる。

○ AI 開発事業者や AI サービス提供事業者が、ウェブサイトが海賊版等の権利侵害複製物を掲載していることを知りながら、当該ウェブサイトから学習データの収集を行ったという事実は、これにより開発された生成 AI により生じる著作権侵害についての規範的な行為主体の認定に当たり、その総合的な考慮の一要素として、当該事業者が規範的な行為主体として侵害の責任を問われる可能性を高めるものと考えられる（AI 開発事業者又は AI サービス提供事業者の行為主体性について、後掲(2)キも参照）。

この記載は「生成・利用段階においては」や「これにより開発された生成 AI により生じる著作権侵害についての」とあることから明らかなように、先ほどの①の問題、すなわち AI 開発事業者における海賊版の収集・利用行為が著作権侵害に該当するという記述ではない⁽³¹⁾。

あくまで、生成・利用段階において学習に用いられた海賊版と同一・類似の AI 生成物が生成された場合に、AI 開発者又は AI サービス提供者が当該生成・利用について行為主体として責任を問われる場合がある、つまり先ほどの②に関する記述である。

また、「考え方」28 頁には以下の記載がある。

○ この点に関して、こうした海賊版等の権利侵害複製物を掲載するウェブサイトからの学習データの収集は、少量の学習データを用いて、学習データに含まれる著作物の創作的表現の影響を強く受けた生成物が出力されるような追加的な学習を行うことを目的として行われる場合もあると考えられる。このような追加的な学習を行うことを目的として、学習データの収集のため既存の著作物の複製等を行う場合、開発・学習段階においては上記イ(イ)のとおり、具体的事案に応じて、学習データの著作物の創作的表現を直接感得できる生成物を出力することが目的であると評価される場合は、享受目的が併存すると考えられるが、これに加えて、生成・利用段階においては、これにより追加的な学習を経た生成 AI が、当該既存の著作物の創作的表現を含む生成物を生成することによる、著作権侵害の結果発生の蓋然性が認められる場合があると考えられる。

○ そのため、海賊版等の権利侵害複製物を掲載するウェブサイトからの学習データの収集を行う場合等に、事業者において、このような、少量の学習データに含まれる著作物の創作的表現の影響を強く受けた生成物が出力されるような追加的な学習を行う目的を有していたと評価され、当該生成 AI による著作権侵害の結果発生の蓋然性を認識しながら、かつ、当該結果を回避する措置を講じることが可能であるにもかかわらずこれを講じなかったといえる場合は、当該事業者は著作権侵害の結果発生を回避すべき注意義務を怠ったものとして、当該生成 AI により生じる著作権侵害について規範的な行為主体と

(31) 「考え方」に関するパブコメに寄せられた意見の中には「海賊版を AI 学習のために複製することについて、30 条の 4 柱書但書に該当して許容されないものと整理すべき」という意見が権利者側から複数寄せられているが（パブコメ 256、257、262、277、278）、文化庁はそのような整理をしていない。

して侵害の責任を問われる可能性が高まるものと考えられる。

この記載には2つの内容が含まれている。

一つ目の○の前半部分（「○ この点に関して・・・享受目的が併存すると考えられるが」の部分については、開発・学習行為についての記述である。

要するに、海賊版等を掲載したサイトから、学習データに含まれる著作物の創作的表現の影響を強く受けた生成物を出力させるような追加学習目的で学習データを収集する行為には享受目的が併存し、法第30条の4が適用されない、という内容だが、これは学習対象著作物が海賊版だから法第30条の4が適用されないという話ではなく、学習目的が表現出力目的だから法第30条の4が適用されないと説明しているにすぎない。

その後引き続き部分（「これに加えて・・・」と二つ目の○の部分）は、やはり、生成・利用段階において学習に用いられた海賊版と同一・類似の AI 生成物が生成された場合に、AI 開発者又は AI サービス提供者が当該生成・利用について行為主体として責任を問われる場合がある、つまり先ほどの②に関する記述である。

（3） 学習禁止意思が付されている著作物

たとえば「AI 学習禁止」など、著作権者の学習禁止意思が付されている著作物があるが、権利制限規定の趣旨及び法第30条の4の立法趣旨からすると、そのような意思表示があることによって法第30条4の適用がない（あるいは法第30条の4の但書に該当する）と解釈することはできない（「考え方」25頁～26頁）。また、そのような一方的な意思表示だけで、著作権者と AI 学習者との間で何らかの契約が締結されたと解釈することもできない。

したがって、著作権者の学習禁止意思が付されている著作物を AI 学習のために利用したとしても著作権侵害には該当しないし、また、学習禁止意思が付されているからといって法第30条の4柱書但書に該当する可能性が高まるわけでもない⁽³²⁾。ただし、著作権者と AI 学習者との間で「対象データを AI 学習に利用しない」「利用する場合には所定の対価を支払う」という契約が真正に成立している場合は、著作権侵害とは別の論点として、当該契約違反になるかという問題は別途生じる⁽³³⁾。

（4） 学習を防止するための機械可読方法による技術的な措置が付されている著作物

① robots.txt やペイウォールのような技術的防止措置

「考え方」26頁では、「AI 学習のための著作物の複製等を防止するための、機械可読な方法による技術的な措置」として、以下のような措置を紹介している。

（例）ウェブサイト内のファイル "robots.txt" への記述によって、AI 学習のための複製を行うクローラによるウェブサイト内へのアクセスを制限する措置

（例）ID・パスワード等を用いた認証によって、AI 学習のための複製を行うクローラによるウェブサイト内へのアクセスを制限する措置

諸外国の立法例では、「オンラインで公衆に利用可能とされるコンテンツのため機械により読み取り可能

(32) 前掲注4・パブコメ207参照

(33) 諸外国では、このような契約を定めても法的強制力を持たないことを明文で定めている例もある（前掲注6・46頁、56頁。英国CDPA第29A条第5項、欧州指令第7条第1項、シンガポール著作権法第187条第1項）。

となる手段のような適切な方法で、同項にいう著作物や他の保護対象物の使用を明示的に「留保」している場合は権利制限の対象とならないとするものがある（欧州 DSM 指令 4 条、ただし、研究組織又は文化遺産機関による学術研究目的の情報解析については対象外）。

日本の著作権法第 30 条の 4 にはこのような限定はないことから、上記のような技術的防止措置を回避して行われる情報解析も権利制限の対象となる⁽³⁴⁾、⁽³⁵⁾。

「(1) 情報解析に活用できる形で整理したデータベースの著作物」のところで説明したように、技術的防止措置が付されていることが「情報解析に活用できる形で整理したデータベースの著作物が将来販売される予定があること」を推認させる一事情にはなり得るが、当該技術的防止措置の回避自体が法第 30 条の 4 柱書但書に該当するわけではない。

② コピーコントロール / アクセスコントロール

技術的防止措置のうち、コピーコントロール / アクセスコントロールのような「技術的保護手段」（著作権法第 2 条第 1 項第 20 号）や「技術的利用制限手段」（同 21 号）に該当するものについては、「技術的利用制限手段」の回避（例：マジコンを用いたゲームのプレイ）については、原則として著作権侵害となる（法第 113 条第 6 項）。

もともと、「技術的利用制限手段に係る研究又は技術の開発の目的上正当な範囲内で行われる場合その他著作権者等の利益を不当に害しない場合」は適用外とされているため、情報解析のために行われる回避行為は基本的にこれに当たる場合が多いと思われる⁽³⁶⁾。

なお、①のような技術的措置は、少なくとも現行法上の「技術的保護手段」や「技術的利用制限手段」には該当しないことから、前述のようにそれを回避したからといって著作権侵害になることはない。

「考え方」27 頁では、以下のような記載がされており、①の技術的措置が現行法上の「技術的保護手段」又は「技術的利用制限手段」には該当しないことを明記している⁽³⁷⁾。

(34) 前掲注 6・66 頁、266 頁

(35) 文化審議会著作権分科会法制度小委員会（第 4 回）早稲田委員発言「これも非常に難しいところではありますが、権利制限規定を技術的な措置で適用がないようにするという、それ自体は権利制限規定が強行法規、強行規定でなくて任意規定というように、解釈されるのだと思いますので、それはいいと思うんですけども、さらにそれを回避して複製した場合はどうなのかというのは、これはなかなか難しい問題ではあるとは思いますが、たとえば 47 条の 5 のインターネット検索の URL を提示するようなどころでは、それなりにそういう技術については、それを回避してはいけないような規定になっておりますので、そういう規定がない限りはちょっとこれもただし書には該当しないんじゃないかなと個人的には考えております。」、文化審議会著作権分科会法制度小委員会（第 4 回）澤田委員発言「(前略) ③の技術的な措置の回避については、これまでの著作権法の中でもたとえば 30 条の私的複製の例外や 47 条の 5 の robot.txt の例で、回避したら権利制限の対象外という規定がわざわざ設けられています。④の海賊版に関しても、30 条や 47 条の 5 の 1 項のただし書で違法なものを用いるケースは権利制限の対象外ということは明記されているところです。そのため、法体系全体の整合性からすると、特にそういった明記のない 30 条の 4 については、③、④の事情があるという一事をもってただし書に当たらないということにはならないのではないかと考えております。」(発言の文脈からすると「当たらない」ではなく「当たる」の誤記ではないかと思われる)。

(36) 前掲注 6・67 頁、前掲注 27

(37) 前掲注 4・パブコメ 213 では「著作権法第 30 条の 4 ただし書きへの該当しうるケースとして、ゲームソフト等に施されている技術的手段に関しては特段の検討がなされていないことに鑑みるに、従来より技術的保護 / 利用制限手段に該当すると考えられてきたゲームソフト等に施されている技術的手段に関しては、それを回避等して行われる複製等が本条によって直ちに制限されるものではなく、ただし書きに該当する可能性が極めて高いと評価されているものと思料。本素案において、これを明記していただくとともに、現時点では技術的保護 / 利用制限手段に該当するかどうかは判然としない技術的手段であっても、権利者の意思を尊重し、著作物に施される技術的な制限を超えて、学習データとして収集されることのないよう、また、AI に活用されることの是非を明確に権利者が意思表示できるよう、当該技術ができる限り技術的保護 / 利用制限手段と評価されることを期待し、技術的保護 / 利用制限手段に関する議論を進めていただくよう要望する。」という意見に対して「法第 30 条の 4 の適用の有無と、技術的保護手段又は技術的利用制限手段該当性とは別個の問題であると考えられます。技術的保護手段又は技術的利用制限手段が施されている場合に、情報解析に活用できる形で整理されたデータベースの著作物が将来販売される予定があることが一定の蓋然性をもって推認されるか否かは、個別具体的な事案に応じて検討すべきものと考えられます。本考え方では、AI 学習のための複製を防止する技術的な措置が技術的保護手段又は技術的利用制限手段に該当するか否かについては、今後の技術の動向も踏まえ検討すべきものとしてされています。」と回答している。

なお、上記のような技術的な措置が、著作権法に規定する「技術的保護手段」又は「技術的利用制限手段」に該当するか否かは、現時点において行われている技術的な措置が、従来、「技術的保護手段」又は「技術的利用制限手段」に該当すると考えられてきたものとは異なることから、今後の技術の動向も踏まえ検討すべきものと考えられる。

(5) 情報解析用 DB 著作物以外の著作物・非著作物のうちライセンス市場が形成されている（すでにライセンス・販売されている）もの

以上述べたとおり、現行法の下で、法第 30 条の 4 柱書但書が適用され著作権侵害となることが明確な学習対象としては、すでにライセンス・販売されている「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物」（情報解析用 DB 著作物）しかない。

もっとも、情報解析用 DB 著作物以外にも、現実にライセンスや販売されている著作物・非著作物（例：単なる DB）は多数ある。たとえば、著作物や情報解析用 DB 著作物に該当しない、新聞社の過去記事 DB や、印刷用やウェブサイトを利用するためのイラストや画像で構成される DB 著作物などもその一例である。

このような「情報解析用 DB 著作物」には該当しないが、ライセンス市場が形成されている（すでにライセンス・販売されている）著作物・非著作物に関する学習についても、法第 30 条の 4 柱書但書に該当する可能性があるとする説もあるが⁽³⁸⁾、一般的には否定されている⁽³⁹⁾。

筆者もこの点については否定説が妥当だと考える。

情報解析用 DB 著作物に該当しない著作物・非著作物については、本来法第 30 条の 4 によって自由に利用できるため、それらの著作物・非著作物についてライセンス市場が形成されているとしても、それを著作権法で保護すべき（ライセンス違反が著作権侵害を構成すると解釈すべき）ことにはならないからである⁽⁴⁰⁾。

2. 4 開発・学習段階での著作権侵害行為について権利者はどの範囲で差止請求等ができるか

開発・学習段階での著作権侵害行為について権利者はどの範囲で差止請求等ができるか。

開発・学習段階での著作権侵害行為が発生した場合、通常は既存著作物の収集・蓄積・学習行為はすでに完了している。

それを前提として、「考え方」29 頁～30 頁、及び同 30 頁脚注 39 では、権利者による 3 種類の差止請求等の可否について以下のように整理がされているが、現行法の解釈として妥当と考える。

(38) 松田政行編『著作権法コンメンタール別冊 平成 30 年・令和 2 年改正解説』（勁草書房、2022 年）32 頁、前掲注 6『AI と著作権』・座談会・今村先生発言 239 頁

(39) 前掲注 6・27 頁、同座談会 238 頁～242 頁、前田健・前掲注 7・208 頁

(40) 前掲注 4・パブコメ 208 で以下の応答がされていることからすると、文化庁も「情報解析用 DB 著作物」以外の著作物のうち、少なくとも享受目的のライセンス市場が形成されているにすぎない著作物の学習については、法第 30 条の 4 柱書但書に該当しないと考えていると思われる。

【意見】ただし書については、「(1) AI 学習のための著作物の複製等を防止する技術的な措置が講じられており、かつ、(2) 楽曲その他の著作物のライセンス市場が構築され、又は構築される予定があることが推認される場合は、(3) この措置を回避して、当該ウェブサイト内に掲されている多数の楽曲等のデータを収集することにより、当該ウェブサイトから AI 学習のための複製等をする行為」の場合も既存のライセンス市場と衝突しうることについてご検討いただきたい。

【回答】本考え方では、従来から法第 30 条の 4 ただし書の該当例として示している「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物が販売されている場合に、当該データベースを情報解析目的で複製等する場合についての考え方を示しています。この場合については、享受目的の販売（ライセンス）市場があることではなく、情報解析目的での販売がされていることが必要と考えられます。

① 将来の AI 学習に用いられる学習用データセットからの除去の請求

著作権侵害の対象となった当該著作物が、将来において AI 学習に用いられることに伴って、複製等の侵害行為が新たに生じる蓋然性が高いといえる場合は、当該 AI 学習に用いられる学習用データセットからの当該著作物の除去が、将来の侵害行為の予防に必要な措置の請求として認められ得る。

② 学習済みモデルの廃棄請求

通常、AI 学習により作成された学習済みモデルは、学習に用いられた著作物の複製物とはいえ、かつ AI 学習により作成された学習済みモデルは、学習データである著作物と類似しないものを生成することができる。

したがって、学習済みモデルについての廃棄請求は、通常は認められない。

もっとも、表現出力目的で学習するなど特殊な学習が行われることによって当該学習済みモデルが、学習データである著作物と類似性のある生成物を高確率で生成する状態にある等の場合は、当該モデルの廃棄請求が認められる場合もあり得る。

③ 学習に用いられた特定の著作物による学習済みモデルへの影響を取り除く措置の請求（例えば特定の学習データを学習用データセットから除去した状態で再度学習済みモデルの作成を行うなど）

現時点では、その技術的な実現可能性や、技術的に可能としてもこれに要する時間的・費用的負担の重さ等（例えば特定の学習データを学習用データセットから除去した状態で再度学習済みモデルの作成を行う場合、当初の学習と同程度の時間的・費用的負担が生じると考えられる。）から、通常、このような措置の請求は認められないと考えられる⁽⁴¹⁾。

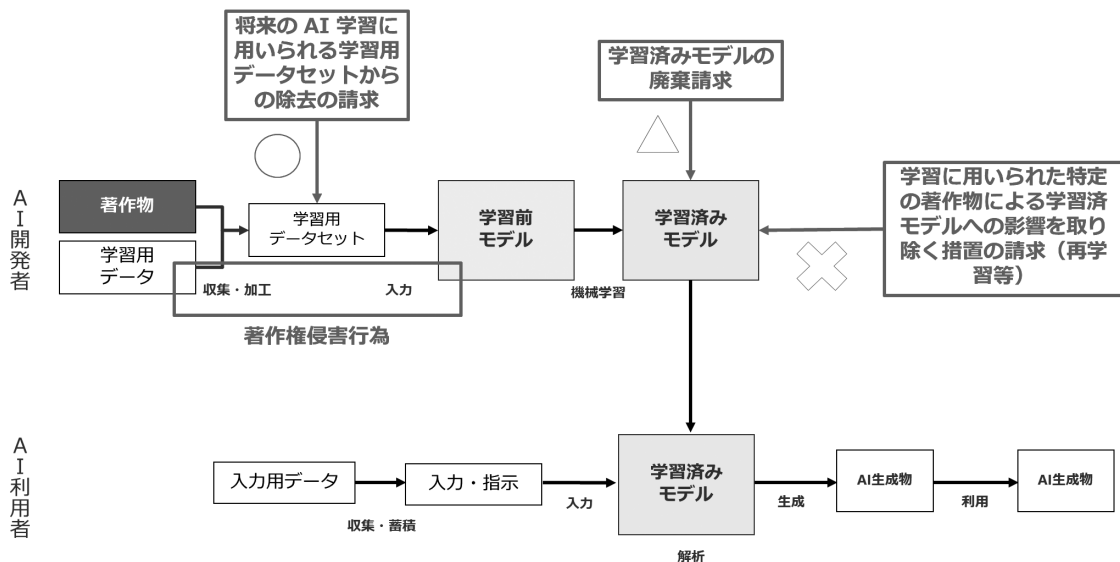


図 16

差止請求等は、ある特定の著作権侵害行為の差止請求であるから、当該著作権侵害に関係のない行為を差し止めることは当然できない。

したがって、仮に、あるサービスにおいて、AI モデル開発・学習段階での著作権侵害行為が発生したとしても、原則として当該サービス全体の差止請求はできないものと思われる。たとえば、ロクラク II 事件最

(41) 「考え方」 41 頁

判（最判平成 23 年 1 月 20 日民集 65 卷 1 号 399 頁〔ロクラクⅡ事件〕）においても、原告が求めた差止請求の具体的な内容は「被告らが著作権あるいは著作隣接権を有する個々の番組を、被告が運営するサービスにおいて複製の対象としてはならない」というものである。

2. 5 生成・利用段階における情報解析と法第 30 条の 4

ここまで、「開発・学習段階」における著作物の利用行為（学習行為）について法第 30 条の 4 が適用されるかという視点から、「学習目的による制限」と「学習対象による制限」について検討してきたが、ここまでの検討結果は、生成・利用段階における既存著作物の入力行為についても、「入力目的による制限」「入力対象による制限」として、基本的にそのまま当てはまる。

(1) 入力目的による制限

「学習目的による制限」と同様、既存著作物の表現出力目的で既存著作物を生成 AI に入力する行為や、同入力目的で蓄積する行為については、法第 30 条の 4 は適用されない（ただし法第 47 条の 5 第 2 項は適用される余地がある）（「考え方」22 頁、37 頁）。

具体的には、画像生成 AI における i2i において既存著作物の加工物を出力する目的で入力する行為や、RAG において既存著作物の創作的表現の全部又は一部を、生成 AI を用いて出力させることを目的として既存著作物をベクトル DB 化する行為（「考え方」20 頁、22 頁）などである。

そして、開発・学習段階においては、学習対象著作物と類似する AI 生成物が生成されるという事実のみから必ずしも表現出力目的が推認されるとは限らないが（「考え方」21 頁）、生成・利用段階においては、開発・学習段階より AI 生成物の生成に、より「近い」段階での行為であるため、既存著作物の入力行為に表現出力目的が推認されることが、より多くなると思われる⁽⁴²⁾。

(2) 入力対象による制限

「学習対象による制限」と同様、「大量の情報を容易に情報解析に活用できる形で整理したデータベースの著作物」（「情報解析用 DB 著作物」）を生成 AI に入力する行為や入力する目的で蓄積する行為については、法第 30 条の 4 但書が適用され、他の権利制限規定の適用がない限り著作権侵害となる。

一方、「海賊版等の権利侵害複製物」「生成 AI への入力禁止意思が付されている著作物」「生成 AI への入力を防止するための機械可読方法による技術的防止措置が付されている著作物」「情報解析用 DB 著作物以外の著作物のうちライセンス市場が形成されているもの」を生成 AI に入力する行為については、開発・学習段階と同様、法第 30 条の 4 が適用され、著作権侵害には該当しない。

2. 6 法第 30 条の 4 と法第 47 条の 5 の役割分担

(1) 情報解析に関する著作権法上の規定

著作権法上、AI の「開発・学習」や「AI 生成物の生成・利用」のような「情報解析」に関する権利制限規定には、法第 30 条の 4 に加えて法第 47 条の 5 が存在する。

この 2 つの権利制限規定は、2018 年の著作権法改正で同時に導入されたが、両者の役割分担は、一言で言うと「情報解析の対象となった入力著作物の類似物を出力する目的がない場合が法第 30 条の 4」「情報解析の対象となった入力著作物の類似物を出力する目的がある場合が法第 47 条の 5」というものである⁽⁴³⁾。

(42) 金子・前掲注 20・57 頁

(43) 前掲注 15

① 法第 30 条の 4 の適用場面

法第 30 条の 4 が適用される典型的な場面は、開発・学習段階において、下図 17 のように開発・学習の対象となった著作物の創作的表現と類似する出力の目的がない場合である。

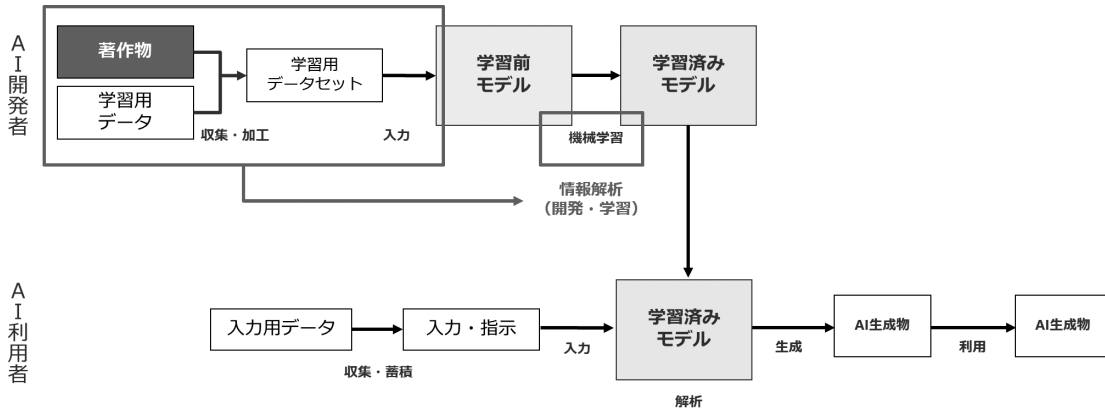


図 17

また、生成・利用段階において、下図 18 のように、入力対象となった著作物の創作的表現と類似する出力の目的がない場合にも法第 30 条の 4 は適用される（「考え方」22 頁、37 頁）。

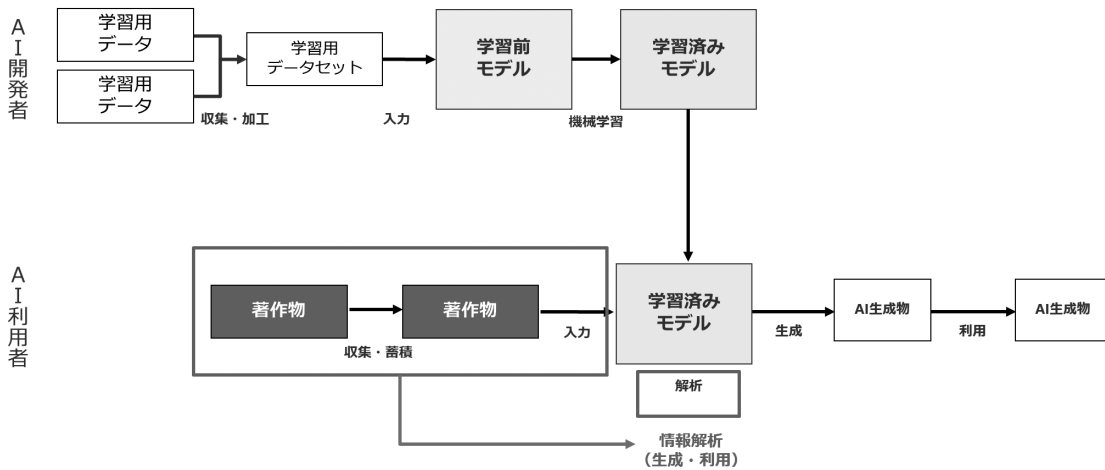


図 18

② 法第 47 条の 5 の適用場面

一方、情報解析の対象となった著作物の類似物を出力する目的がある場合が法第 47 条の 5 の適用場面であるが、これは 2 つのパターンに分かれる。

1 つは、開発・学習段階において、開発・学習段階における情報解析目的に加えて、下図 19 のように開発・学習の対象となった著作物の創作的表現と類似する出力を行う目的がある場合（享受目的併存）である。

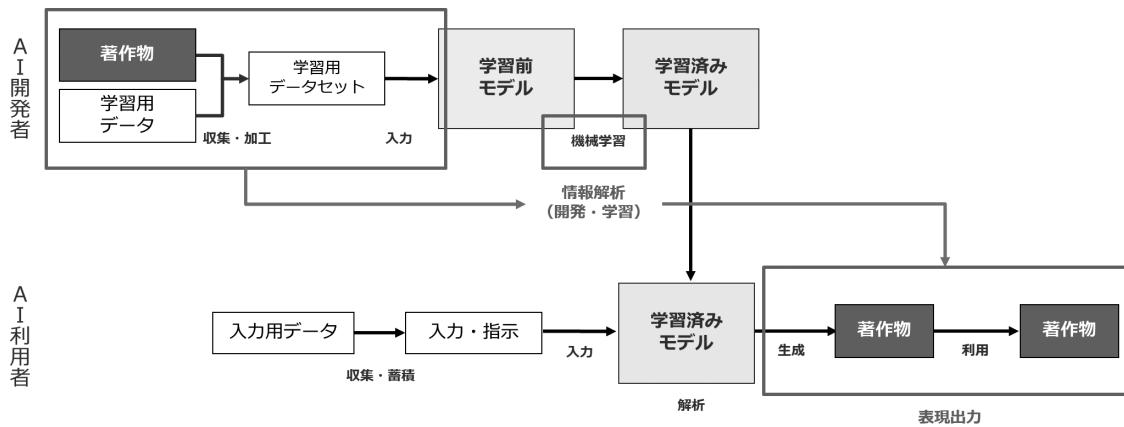


図 19

このパターンにおいては、著作物の利用行為は2種類ある。

1つは、情報解析⁽⁴⁴⁾の結果としての著作物の利用行為、もう一つは、そのような著作物の利用行為の準備のために行われる著作物の利用行為である。前者に関する権利制限規定としては法第47条の5第1項第2号が、後者については法第47条の5第2項が適用される(図20、ただし同条の要件を満たすかはまた別の問題である)。

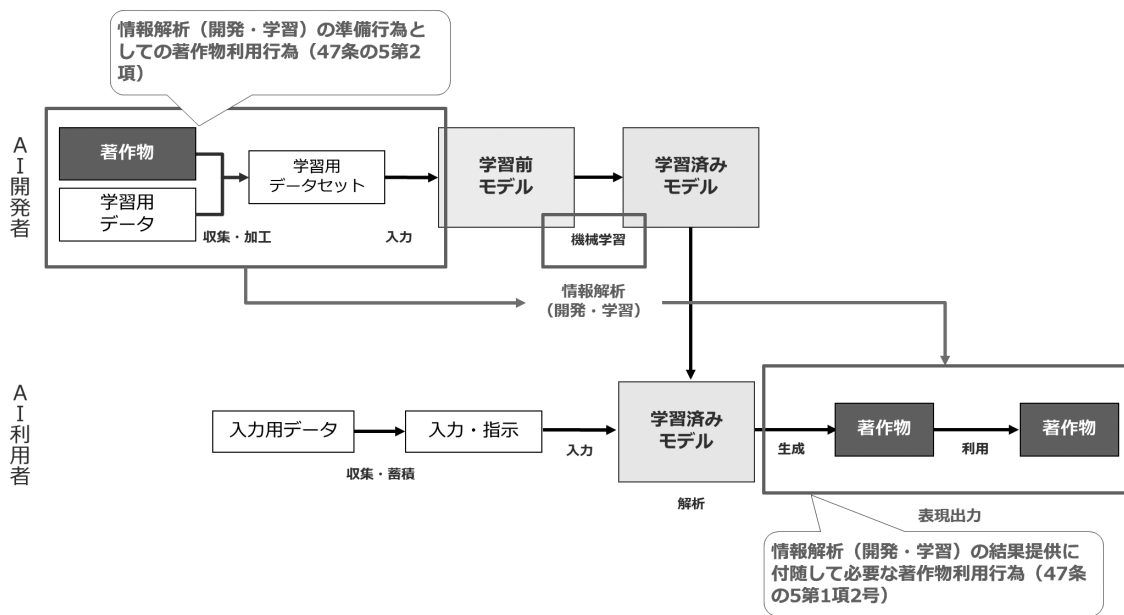


図 20

もう一つのパターンは、下図 21 のように、生成・利用段階において、入力対象となった著作物の情報解析目的に加えて、入力(情報解析)の対象となった著作物の創作的表現と類似する出力の目的がある場合である。

(44) このパターンにおいて、法第47条の5の「情報解析」に該当する行為は2つある。1つは、開発・学習段階における学習済みモデルの「学習」であり、もう一つは、生成・利用段階における、学習済みモデル内での入力データの「解析」である。最終的に出力されている AI 生成物は、この両方の「情報解析」の結果生成されているといえるため、当該 AI 生成物の生成・出力行為は、これら2つの「情報解析」の結果としての著作物の利用行為といえる。

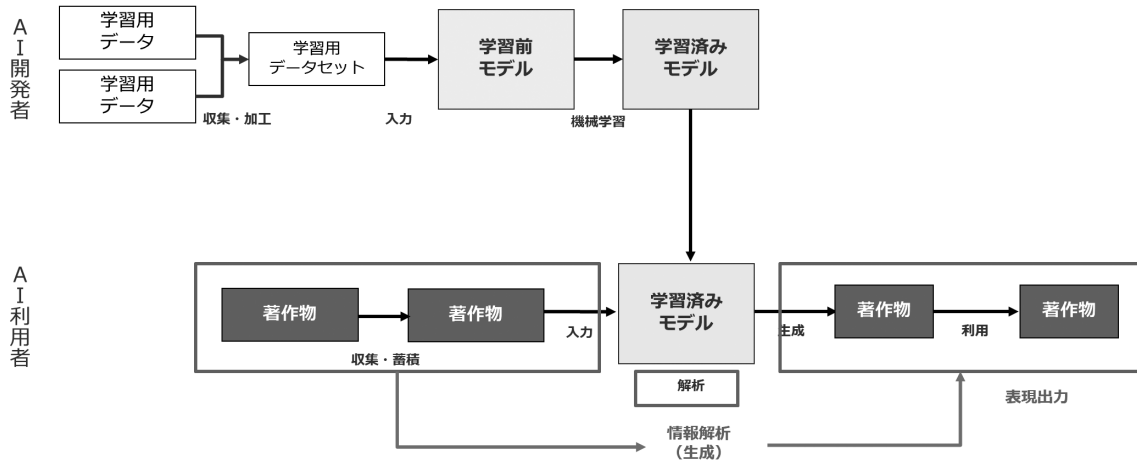


図 21

このパターンにおいても、先ほどの図 19、20 と同様、情報解析の結果としての著作物の利用行為及び著作物の利用行為の準備のために行われる著作物の利用行為が行われ、それぞれについて権利制限規定として法第 47 条の 5 第 1 項第 2 号、同 2 項が適用されうる (図 22)。

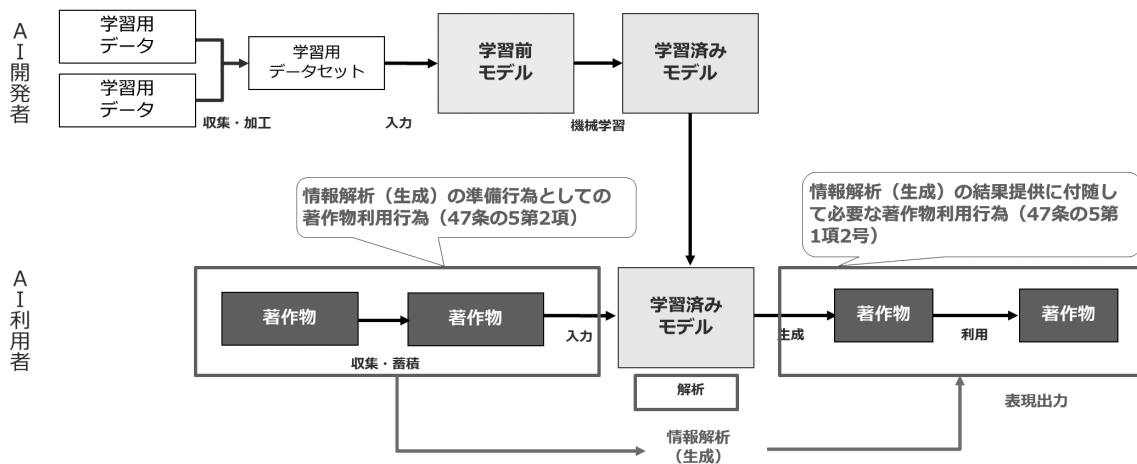


図 22

(2) 実際の検討の順番

法第 30 条の 4 と法第 47 条の 5 の役割分担は上記のとおりだが、法第 30 条の 4 が適用される場面は著作物の非享受目的利用であるのに対して、法第 47 条の 5 が適用される場面は著作物の (軽微な) 享受目的利用であるため、後者の方が適用要件が格段に厳しい。

そのため、「情報解析」のための既存著作物の利用可否を検討する場合には、まずは法第 30 条の 4 の適用可否を検討し、同条が適用されない場合には、法第 47 条の 5 の適用可否を検討するという順番で検討することになる。

下図 23 が検討のフローチャートである。

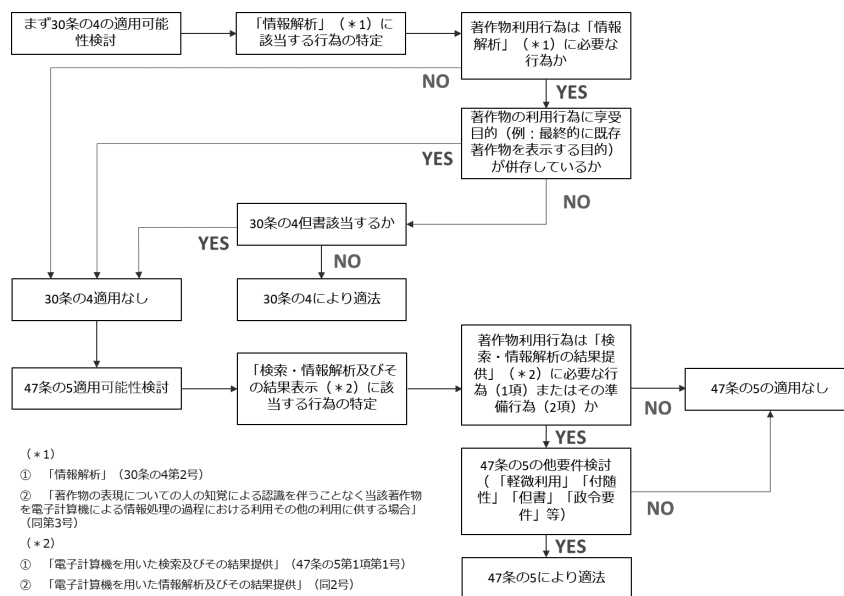


図 23

(3) 法第 47 条の 5 は使い道があるのか

もともと、法第 30 条の 4 が適用できない場合に法第 47 条の 5 が適用可能だとしても、特に画像生成 AI のようなコンテンツ生成 AI の場合は、法第 47 条の 5 が適用されて適法となる可能性は非常に低いと思われる。これは、画像生成 AI については、特に法第 47 条の 5 の要件のうちの「付随性」を満たすことが困難だからである。

法第 47 条の 5 第 1 項において、検索又は情報解析の対象となった既存著作物の一部を表示する行為（「著作物利用行為」）は、検索又は情報解析の結果提供行為（「結果提供行為」）に「付随して」の要件を満たす必要がある。

このような「付随して」の要件が要求されるのは、所在検索サービス・情報解析サービスの主たる目的はあくまで結果提供行為にあり、著作物利用行為は、当該「結果」が当該サービスのユーザーの求める情報であるか否か容易に確認することができるようにするために限定的に許容されているにすぎないためである。

「付随して」と言えるためには、①生成文及び書誌情報（「結果」）の提供行為（結果提供行為）と、当該提供に伴う情報解析又は検索対象となった既存著作物の一部を表示する行為（著作物利用行為）の区分可能性及び②前者が「主」、後者が「従」の関係になければならない⁽⁴⁵⁾。

結果提供行為と著作物利用行為が一体化している場合については、区分可能性がなく、また著作物の利用が主たる目的であることも多いと考えられるため「付随して」の要件を満たさない⁽⁴⁶⁾。

これを前提とすると、画像生成 AI のようなコンテンツ生成 AI の出力について法第 47 条の 5 が適用される可能性は非常に低いと思われる（「考え方」22 頁脚注 25）⁽⁴⁷⁾。

これは、情報解析結果（AI 生成物）に、情報解析対象著作物（学習に利用された著作物や入力に利用された著作物）の創作的表現が含まれている場合、生成された AI 生成物の中に、元の著作物の創作的表現と情報解析の結果が渾然一体となってしまっているためである。

一方、テキスト生成 AI（たとえば RAG）の場合は、見せ方次第では付随性・軽微性を満たすことも十分考えられるため、法第 47 条の 5 の適用を検討する意義は大きいと思われる（「考え方」22 頁）⁽⁴⁸⁾。

以上

(45) 前掲注 38・103 頁

(46) 前掲注 38・45・104 頁

(47) 前掲注 6・22 頁、同座談会・242 頁～247 頁

(48) 前掲注 6・24 頁、前掲注 27