HOKKAIDO UNIVERSITY

Title	生成AIと著作権 ： 「考え方について」に関する3つの論点
Author(s)	奥邨, 弘司
Citation	知的財産法政策学研究, 70, 103-143
Issue Date	2025-06
Doc URL	https://hdl.handle.net/2115/95534
Type	departmental bulletin paper
File Information	70-03_Okumura.pdf



Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

生成AIと著作権
～「考え方について」に関する3つの論点～

奥　邨　弘　司

Ⅰ　はじめに

今日は、生成 AI と著作権の問題を考える上で、必読の資料である「AIと著作権に関する考え方について[1]」（以下、「考え方について」）の中で、特に議論の多い 3 つの論点に絞って、かなり細かいところをお話ししたいと思っております。具体的には「享受目的の併存について」「30 条柱書但書について」「依拠について」の 3 つとなります。

なお、今日は、「考え方について」を一通り読んでおられることを前提にお話しすることになりますので、その点、ご容赦ください。

Ⅱ　享受目的の併存について

3 つの論点の第 1 は、享受目的が併存する場合には、30 条の 4 が適用されないという解釈についてです。

論点に入る前に、前提となる議論を整理しておきましょう。

1　機械学習に伴う利用行為に関する権利制限規定

改めて条文を確認すると、30 条の 4 柱書は、利用される「著作物に表

＊本稿は、2024 年 8 月 29 日に行われた北海道大学サマーセミナーにおける筆者の講演内容に、加筆修正を加えたものである。

[1] 文化審議会著作権分科会法制度小委員会「AI と著作権に関する考え方について令和 6 年 3 月 15 日」。

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」には、必要な限度で著作物を利用できると規定しています。そしてこの「享受・・・を目的としない場合」つまり非享受目的の場合の例として、1 号から 3 号が定められており、その中の 2 号に「情報解析[2]・・・の用に供する場合」があるわけです。さらに、AI を訓練する際の機械学習が、この情報解析に含まれると一般的には解釈されています[3]。結果、著作物の利用目的が機械学習だけの場合、それは非享受目的の利用だから、権利者の許諾なく、機械学習に伴う著作物の利用(主に複製)が可能になると解するわけです。

この 30 条の 4 は、柔軟な権利制限規定の 1 つとして、平成 30 年改正で導入された規定[4]ですが、生成 AI が登場するまでは、ことさら世間一般の関心を集めることはありませんでした。しかし、生成 AI が注目される中で、一躍、世間の関心は 30 条の 4 に向かい、結果、喧々諤々の議論が始まったわけです。

しかし、平成 30 年改正に向けた審議会での議論に参加していた身としては、生成 AI に関する機械学習(に伴う複製など)との関係で、世間が 30 条の 4 ばかりを議論をするのは、大いに不思議な気がしていました。と言うのも、機械学習、正確には機械学習を含む情報解析に関する権利制限規定としては、30 条の 4 だけではなくて、47 条の 5 も存在するからです。本来なら、30 条の 4 と 47 条の 5 の両方について議論しなければいけないのに、なぜか、後者については、ほとんど議論されることなく、みんなで前者だけに集中して議論してしまっているんです。私は、この点は、生成 AI と著作権に関する議論において、ある意味不幸なことだったと思っております。何が不幸だったのかは、後ほど触れたいと思います。

---

[2] 30 条の 4 括弧書きによれば、情報解析とは「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、影像その他の要素に係る情報を抽出し、比較、分類その他の解析を行うこと」を言う。

[3] 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)令和元年 10 月 24 日」(以下、基本的な考え方)10 頁参照。

[4] 正確には、旧 30 条の 4 を全面的に改正した規定。

２　47 条の 5 について

まず、47 条の 5 について見ておきましょう。

47 条の 5 ですが、かなり長い規定ですので、条文を引用するのは控えたいと思います。思い切り単純化して説明すると、47 条の 5 は、著作物を、①情報検索[5]や情報解析[6]のために複製など[7]すること、そしてそのようにして複製した著作物を、②情報解析などの結果を提供する際に、ちょっとだけ見せたり聞かせたりなどすること[8]を可能とする規定です。

インターネット検索サービスを例にすると、検索のために、webサイトの情報をダウンロードすること、ダウンロードしたものを検索しやすいように加工する際に複製すること、検索の過程でデータを複製することが前記①であり、検索結果として URL を提供する際に、複製した web サイトの情報の一部をスニペットやサムネイルとして表示することが前記②であり、それらが 47 条の 5 によって可能となるわけです。

47 条の 5 には、1 項と 2 項がありますが、検索結果などの提供に付随してスニペットなどとして著作物の一部を（インターネットで送信して）表示などすること（前記②相当。以下、付随的軽微利用）を可能とするのが 1 項、付随的軽微利用に至るまでの複製行為、例えば、検索サービスなら、web サイトのダウンロードや各種データ加工などに伴う複製（前記①相当。以下、まとめて準備行為のための複製）を可能とするのが 2 項になります[9]。物事の順番としては、準備行為のための複製があって、その後、付随的軽微利用が行われるわけですが、条文上は、前者が 2 項、後者が 1 項となっており、少し分かりづらくなっております。

ここで注目していただきたいのは、付随的軽微利用というのは、軽微ではあっても著作物を見せたり聞かせたりして享受させますから、享受目的の利用になります。そして、準備行為のための複製は、もちろん、情報検索や情報解析などのような非享受を目的とした複製ですが、同時

---

[5] 47 条の 5 第 1 項 1 号。

[6] 同項 2 号。

[7] 正確には、複製、公衆送信、頒布。

[8] 正確には、いずれの方法での利用も許される。

[9] 「基本的な考え方」前掲注 3）22～24 頁参照。

に、後の付随的軽微利用をも目的として行われるわけですので、そこには、享受目的が同時併存しているわけです。

整理しますと、30 条の 4 は非享受目的利用に関する権利制限規定でしたが、47 条の 5 は非享受目的と享受目的が同時併存する場合の権利制限規定ということになります。平成 30 年改正の際には、そういう区分で、2 つの条文を作ったわけです。

３　30 条の 4 と 47 条の 5 の使い分け

では、機械学習に関して、30 条の 4 と 47 条の 5 をどう使い分けるかを、いくつかの例をあげて見ていきたいと思います。

(1) 識別系ＡＩの場合

最初は、識別系の AI の場合です。

例えば、ピカソ作かもしれないという絵の真贋を、パーセントで表示する AI を開発するとします。このような AI を開発するためには、本物のピカソの絵をたくさん機械学習させないといけません。その際には当然、ピカソの絵がたくさん複製されます。

この AI は、真贋結果をパーセント表示するだけで、学習対象となったピカソの絵自体を出力して人間に見せることはありませんので、ピカソの絵の複製に際して、享受目的はありません。あくまでも機械学習＝情報解析する目的、つまり非享受目的だけですので、この場合の複製については、30 条の 4 の守備範囲となります。

ところで、真贋をパーセント表示するだけだと、AI の判定結果が正しいのかどうか、人間としては不安になります。そこで、結果をパーセントで表示するだけではなくて、例えば、判定対象の絵と一番よく似たピカソの絵をサムネイルで見せて、その絵との異同を簡単に説明することで、パーセント表示の補足をするような AI を開発することにしたとします。

この AI の場合は、ピカソの絵を、機械学習目的(＝非享受目的)とサムネイルで見せる目的(＝享受目的)の両方の目的で複製しているわけです。つまり、ピカソの絵を複製する時点で、非享受目的と享受目的が同時に

併存しているわけです。この場合、非享受目的だけではないので当然 30 条の 4 は適用されません。こういう非享受目的と享受目的の併存型の機械学習に伴う複製は、47 条の 5 の守備範囲となります。

47 条の 5 で許される享受目的利用は、付随的軽微利用に限られますが、この場合、サムネイルの表示は、真贋判定結果のパーセント表示と真作との異同の簡単な説明という、情報解析の結果の提供に付随していますし、サムネイルですから軽微でもありますので、付随的軽微利用といえるでしょう。なお、付随的軽微利用に関する部分は 47 条の 5 第 1 項に基づき、付随的軽微利用の準備に関する部分つまり準備行為のための複製、ここには機械学習に伴う複製も含まれますが、それについては同条 2 項に基づくことになります。

(2) 生成ＡＩの場合

ここまで説明したのは識別系の AI でした。では、生成 AI の場合は、30 条の 4 と 47 条の 5 をどう使い分ければ良いでしょう。

例えば、キュービズムに共通する画風を理解して、その画風で絵画を描く AI を開発する過程で、キュービズムの絵画を機械学習に伴って複製する場合を考えましょう。この場合、絵画の複製目的は、キュービズムに共通する画風を機械学習＝情報解析することだけ、つまり、非享受目的だけなので、30 条の 4 が適用されるケースになります。

ここで、それは違うのではないか、と思われた方もいるかもしれません。と言うのも、この AI は生成 AI ですから、画像(ここではキュービズムの画風の画像)を出力します。とすると、人間に見せて享受させる目的がある画像を出力するのだから、この場合の機械学習に伴う複製には、享受目的が併存しているのではないだろうか。識別系で言えば、サムネイルを出力するタイプと同様に考えるべきであり、30 条の 4 ではなくて 47 条の 5 の領分になるのではないか。そう思われた方もいるのではないでしょうか。

結論から言えば、そのような解釈は正しくありません。

理由は、まさに、サムネイルを出力するタイプと比べて頂くと分かります。サムネイルを出力するタイプの場合、出力されるのは、学習対象の絵画のサムネイル、イコール、縮小複製物です。複製物ですから、著

作権法上は、同じ著作物となります。つまり、学習対象著作物自体が――縮小されているとは言え――出力され、結果、鑑賞＝享受の対象となっているわけです。

では、キュービズムの画風を再現する生成 AI の場合はどうでしょうか。生成され出力される絵画は、学習された絵画ではありません。あくまでも、AI が生成した別の絵画です。確かに、学習対象画像の画風と出力画像の画風は、いずれもキュービズムで共通しています。しかし、画風はアイデアであって[10]、著作権では保護されません。著作権が保護するのは、あくまでも創作的表現です。

2 つの絵画があって、両者で共通するのがアイデアだけで、創作的表現が共通しない場合、その2つの絵画は著作権法上、別物と評価されます。つまり、この場合、キュービズムの画風を再現する生成 AI は、学習した著作物とは別物を出力して享受させているに過ぎません。この点が、サムネイルを出力する AI との大きな違いです。

ここで、30 条の 4 柱書を改めて見てみましょう。すると、「著作物は・・・当該著作物に表現された思想又は感情を・・・享受・・・(す)ることを目的としない場合には・・・利用することができる」とあります。下線を引いた部分に注目してください。享受の対象は「当該著作物」です。そして、この「当該著作物」が何を指すかと言うと、利用される著作物、ここで言えば、学習に伴い複製される著作物＝学習対象著作物を指します。つまり、30 条の 4 を適用するためには、学習対象著作物を享受させる目的がないことが求められているのであって、学習対象著作物とは別のものを享受させる目的があっても、それは 30 条の 4 の適用を排除しません[11]。

ここまでの分析を前提にすると、次のようなことも言えます。

もし、学習対象著作物と、アイデアだけでなく、創作的表現も共通するものを出力する目的がある場合、その場合は、先ほどのサムネイルを

[10] 一般的に画風や絵柄と呼ばれるものが、アイデアなのかそれとも表現なのかについて、衆目が一致する定義等は存在しないことを前提とした上で、説明の便宜のため、本稿における用語法として、画風はアイデアに属し、絵柄は表現に属するものとする。

[11] 「考え方について」前掲注 1）19 頁注 21 参照。

出力する目的がある場合と同様となり、学習に伴う複製の時点で、享受目的が同時併存することになって、30 条の 4 の適用はなくなり、47 条の 5 の領分になるわけです。先の例で言えば、キュービズムの画風を機械学習する目的だけでなくて、学習対象絵画の絵柄（創作的表現）の一部を出力画像に含める目的もあった場合、前者は非享受目的ですが、後者は享受目的なので、両目的の同時併存となって、30 条の 4 の適用を外れてしまうわけです。

なお、この場合、47 条の 5 の適用の余地がありますが、同条で許される享受目的利用は、付随的軽微利用に限られますので、適用範囲は広くはありません。仮に適用される場合は、付随的軽微利用に関する部分は、47 条の 5 第 1 項、付随的軽微利用の準備に関する部分つまり準備行為のための複製（機械学習に伴う複製も含まれる）については同条 2 項に基づくことは、識別系 AI の場合と同様です。

４　単独適用説

ここまで、機械学習に伴う複製に関して、享受目的が同時併存するか否かで、30 条の 4 と 47 条の 5 を使い分ける必要がある旨を説明してきました。これは、「考え方について」に示された考え方でもあります[12]。以下では、仮に、使い分け説と呼びます。

ところで、学説の中には、使い分け説とは別に、単独適用説とでも呼ぶべき考え方があります。具体的には、機械学習には 30 条の 4 のみが適用されるとする説です[13]。

単独適用説は、30 条の 4 の条文構造から生まれます。すなわち、30 条

[12] 「考え方について」前掲注 1）20～22 頁参照。

[13] 文化審議会著作権分科会（2024 年 3 月 19 日）髙部眞規子委員発言（「今の条文は、情報解析というものを 30 条の 4 の第 2 号で、享受し又は享受させることを目的としない場合の例示として挙げています。そのような条文構造からは、情報解析に当たるとしながら享受目的が併存するので 30 条の 4 に当たらないという説明の仕方というのは、ちょっと難しいような気がいたします。」）および愛知靖之「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『AI と著作権』（勁草書房・2024）15～20 頁参照。

の 4 柱書は、「次に掲げる場合その他の・・・享受(する)ことを目的としない場合には・・・(著作物を) 利用することができる」と規定しています。この規定中、「次に掲げる場合」というのは、同条1号、2号、3号を指していますので、これを組み込むと、先の引用部分は「1 号、2 号または 3 号の場合その他の・・・享受(する)ことを目的としない場合には・・・(著作物を) 利用することができる」となります。2 号の「情報解析」だけに注目しますと先の引用部分は「**情報解析の用に供する**場合その他の・・・享受(する)ことを目的としない場合には・・・(著作物を) 利用することができる」と書き換えることができます。

ところで、法律用語の一般的な用法として、「その他の」は例示を意味するとされていますので、「情報解析の用に供する場合」は「享受(する)ことを目的としない場合」の例示に当たるわけです。としますと、「情報解析の用に供する場合」に当てはまれば、自動的に非享受目的となるということになるのではないか、というわけです。

ここで、先ほども申しましたように、機械学習というのは情報解析に含まれると理解されていますので、つまり、「機械学習の用に供する場合」は自動的に非享受目的となり、30 条の 4 が適用され得るという解釈になるわけです[14]。言い方を変えると、30 条の 4 の条文構造上、機械学習の用に供する場合に、享受目的が併存しているということはあり得ないというわけです。

5　単独適用説への疑問

30 条の 4 の条文構造だけに注目して文理解釈をすると、単独適用説のような解釈が導かれるのかもしれません。しかし、単独適用説には、いくつか深刻な問題があります。それを見ていきましょう。

[14] ただし、単独適用説も、使い分け説で享受目的の同時併存を問題とするような状況については、30 条の 4 の「必要と認められる限度」要件や、柱書但書の解釈によって、同条の適用対象外とするようである。例えば、注 13)で引用した髙部委員発言の直後の部分参照。

(1) 文理解釈上の問題

単独適用説は、30 条の 4 の文理解釈を根拠とするわけですが、文理解釈を言うのであれば、30 条の 4 だけではなくて 47 条の 5 についても行わなければならないはずです。

47 条の 5 第 1 項は「電子計算機を用いた情報処理により新たな知見又は情報を創出することによつて著作物の利用の促進に資する次の各号に掲げる行為を行う者・・・は」と定めており、2 号は「電子計算機による情報解析を行い、及びその結果を提供すること」規定しています。つまり下線部を 2 号に置き換えて「電子計算機を用いた情報処理により新たな知見又は情報を創出することによつて著作物の利用の促進に資する**電子計算機による情報解析を行い、及びその結果を提供すること**を行う者・・・は」とすることが可能です。そして、情報解析には、既に見たように機械学習が含まれますし、AI の場合、それらが電子計算機によるものであることも明確で、結果も提供します。つまり、47 条の 5 を普通に文理解釈すると、機械学習に伴う複製について、同条が適用されることは明確です[15]。

単独適用説は、文理解釈を理由に、機械学習に伴う複製について、30 条の 4 の単独適用を主張するわけですが、同じく文理解釈すると、機械学習に伴う複製について、47 条の 5 も適用されるはずです。30 条の 4 の「単独」適用のためには、47 条の 5 の適用を排除しないといけませんが、それはどのように説明するのでしょうか。そこが説明できないと思うのです[16]。

(2) 47 条の 5 の不要化

逆に言うと、単独適用説に立つと、47 条の 5 は不要になります。

先ほど、ピカソ作とされる絵画の真贋判定をする AI のうち、パーセン

---

[15] 正確には、先にⅡ.3.「30 条の 4 と 47 条の 5 の使い分け」で説明したように、機械学習に伴う複製は、準備行為のための複製に含まれるので、47 条の 5 第 2 項が適用される。

[16]「座談会」上野＝奥邨・前掲注 13）209～211 頁（奥邨弘司発言）および前田健「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 Online ロージャーナル（2024 年 1 月 29 日）Ⅱ.1.(3)参照。

ト表示だけでなく、サムネイルも表示するタイプを開発する場合の機械学習に伴う複製は、30 条の 4 ではなくて、47 条の 5 が適用されるとお話ししましたが、これは使い分け説の場合の説明になります。

もし、使い分け説ではなくて、単独適用説を採用すると、この場合も、30 条の 4 の適用になります。なぜなら、単独適用説だと、機械学習＝情報解析＝非享受目的なので、それ以上に享受目的――この場合は、サムネイルを見せる目的――があるかどうかを検討する必要はないからです。同じように、生成 AI で、学習対象著作物の創作的表現の一部を出力に含めることを目的とするようなタイプのものを開発する場合の機械学習に伴う複製についても、30 条の 4 が適用されることになります。

識別系の AI でも、生成 AI でも、いずれにしても、機械学習に伴う複製に関しては、常に 30 条の 4 のみが適用される、まさに単独適用説と呼ぶ所以となります。

ただ、そうしますと、47 条の 5 は不要になります。先ほど見ましたように、30 条の 4 も 47 条の 5 も、両方とも、機械学習を含む情報解析の場合における著作物の利用について、著作権を制限する旨を規定していることが文言上明らかです。しかも、両規定は、平成 30 年改正時に、同時に導入された、いわば兄弟・姉妹規定です。と言うことは、普通に考えますと、機械学習に関する著作物の利用に関して、両条の間で、役割分担があると考えるのが自然なはずです。言葉を換えますと、機械学習に関する著作物の利用について、30 条の 4 が適用されるべき場合と 47 条の 5 が適用されるべき場合が、ちゃんと存在するはずです[17]。

しかしながら、単独適用説に立って、機械学習すなわち情報解析に当たる以上、非享受目的であり、自動的に 30 条の 4 が適用されると解釈すると、47 条の 5(特に 2 項)はなくても良いということになってしまいますが、同時期に導入された規定であるにもかかわらず、そのうちの一方が(一部）不要になるような解釈は一般的と言えるのでしょうか。

逆に言えば、単独適用説に立つと、内閣法制局も、国会も、30 条の 4 があれば、47 条の 5(第 2 項)は不要なのに、そのことを見逃して立法してしまったと解釈するに等しいように思うのですが、果たしてそれで良い

[17] 同上。

のでしょうか。

もちろん、改正から 50 年も 100 年もたって、世の中も裁判例も大きく変化した、そういう状況ならば、同時期の改正で導入された規定の一部をないものとして扱うという解釈もあり得ないわけではないと思います。しかし、30 条の 4 と 47 条の 5 が導入されたのは、わずか 5・6 年前のことです。裁判例だってまだありません。にもかかわらず、一部の規定を無意味化する解釈をとるというのは、私自身はなかなか得心できない部分があるわけです。

(3) 立法経緯との不整合

単独適用説は立法経緯とも整合しません。

平成 30 年改正の前には、30 条の 4 のもとになった条文として、旧 47 条の 7 がありました。旧 47 条の 7 は、統計的情報解析のためなら、必要な範囲で自由に複製ができるとする規定でした。ただ、この規定は解析のための複製を許すだけであり、解析結果を表示するときに、解析時に複製したものの一部を出力することまでは許すものではありませんでした[18]。先のピカソ作とされる絵画の真贋判定 AI を例に説明すると、真作率をパーセント表示する AI の場合は適用があるけれど、学習したもののサムネイル表示まで目指す場合は適用がない規定だったわけです。

しかし、それではあまりにも使い勝手が悪いということで、平成 30 年改正に際して、付随的で軽微ならば、学習対象著作物を一部出力する場合にも適用できる 47 条の 5 が作られたわけです[19]。つまり、立法の経緯に照らせば、現行 30 条の 4 は、もとになった旧 47 条の 7 と同様、学習対象著作物を出力する目的があるときは適用されない一方、その場合に対応できる規定として、47 条の 5 が立法化されたのです。

---

[18] 池村聡「著作権法コンメンタール別冊　平成 21 年改正解説」（勁草書房・2010）118 頁および半田正夫＝松田政行編「著作権法コンメンタール　第 2 版」（勁草書房・2015）578 頁〔奥邨弘司担当〕参照。

[19] 文化審議会著作権分科会　法制・基本問題小委員会　新たな時代のニーズに的確に対応した制度等の整備に関するワーキングチーム（第 2 回）富士通株式会社提出資料（特に 7〜8 頁）参照。あわせて、「座談会」上野＝奥邨・前掲注 13）214 頁（奥邨弘司発言）参照。

このように、両規定の間では、機械学習に関して、役割分担を前提とする立法経緯があるはずなのですが、単独適用説だと、役割分担が出てこず、47条の5(特に2項)がいらなくなってしまうわけです。つまり、単独適用説は立法経緯にも整合しないのです。

(4) 適用範囲の相違

文理解釈の観点、条文間の役割分担の観点、および、立法経緯との整合性の観点から、単独適用説の問題点を指摘しました。いずれも、研究者としては、かなり重要な問題だと考えております。ただ、実務的に考えれば、30条の4と47条の5(特に2項)とで、規定内容に実質的な差がないのであれば、どちらが適用になっても、問題はないのでは、というご意見もあるように思います。

しかしながら、実は、両規定には、結構な差があるのです[20]。この点が、往々にして忘れられてしまっているように思います。例えば、47条の5ですと、権利制限が適用される主体が限定されています[21]。具体的には、政令[22]に指定された条件を守っていることが求められています。また、利用できる著作物についても、47条の5は限定があります[23]。ところが、30条の4には、そういう限定が一切ないんですね[24]。そうすると、同じ機械学習であっても、47条の5が適用されるのか、それとも30条の4が適用されるのかによって、本来は権利制限される範囲に差が出てきます。端的に言えば、30条の4の方が制限される範囲が広い。にもかかわらず、単独適用説を採用すると、その差が全部なしになってしまいます。しかも、より広く権利制限されるんです。これはかなり大きな問題だと思う

---

[20] 「座談会」上野＝奥邨・前掲注13) 211頁(奥邨弘司発言)および前田・前掲注16) Ⅱ.2.(2)参照。

[21] 「基本的な考え方」前掲注3) 20頁参照。

[22] 法47条の5第1項の主体については、著作権法施行令7条の4第1項。法同条2項の主体については、施行令7条の4第2項。

[23] 法47条の5第1項については、公衆への提供または提示(送信可能化を含む)が行われた著作物(以下、公衆提供等著作物)のうち、公表された著作物または送信可能化された著作物に限られ、第2項については、公衆提供等著作物に限られる。

[24] 未公表著作物も対象となる。「基本的な考え方」前掲注3) 40頁参照。

んです。

ここで、30 条の 4 と 47 条の 5 で、機械学習に関する権利制限の範囲に差があるのはなぜかを考えてみましょう。これは、偏に、享受目的の有無に尽きます。具体的には、47 条の 5 の場合、軽微ですけれども、学習対象著作物を見せたり聞かせたりして、享受させる目的があるわけです。その結果、そういう目的のない 30 条の 4 の場合と比べると、権利者に与える影響は、理論的に、そして、類型的に、大きくなる可能性があります。その影響を、できる限り小さくしようという配慮があって、先ほどお話ししたような限定が、47 条の 5 には付されているわけです。

にもかかわらず、単独適用説を採用すると、機械学習に関しては、47 条の 5 が適用される余地がありませんので、そういう配慮がなくなってしまいます。私は、この点は、かなり深刻な問題だと思います。

もしかしたら、但書を活用して、その中で、そういう配慮も行えば良いのでは、というご意見も出そうですが、但書は総合考慮規定でしかありませんので、予測可能性が極めて低くなってしまいます。何より、立法府が明文で規定した要件、しかもわざわざ行政府が政令で明確化した要件をまるっと無視して、漠然とした但書で司法府が判断するというのは、どう考えても穏当な処理だと私には思えません。

(5) 小括

以上が、単独適用説への疑問点となります。また、後ほど触れますが、私としては、30 条の 4 と 47 条の 5 の使い分けを上手く行うことで、オプトアウトや海賊版の学習などの問題に対応する余地が拡がると思っています。その辺は、残念ながら、単独適用説では無理であり、その点でも、単独適用説はお薦めできません。

最後に、2 つ補足です。1 点目は、使い分け説は機械学習が非享受目的であることに疑問を呈しているわけではないという点です。

一部誤解があるんですが、機械学習に、非享受目的だけのものと、享受目的と非享受目的が併存しているものの 2 種類があると考えているわけではありません[25]。使い分け説でも、機械学習は非享受目的です。ただ、

[25] なお、30 条の 4 の場合の情報解析は非享受目的だけのものを指し、47 条の 5 の

1つの行為には、目的が1つの場合もあれば、複数ある場合もあることを前提とした上で、学習対象著作物の複製の目的が機械学習だけか、それとあわせて他に(享受)目的があるのかを問うているのが、使い分け説なのだと私は思います。

2点目は立証責任との関係です。使い分け説については、同説の場合、30条の4が適用されるためには、2号の「情報解析」目的に該当することを示すだけではダメで、それ以外に、享受目的がないことも示すことが求められるところ、「ない」ことの証明は難しく、同条の適用を求める者にとって、負担が大きいのではないか、また、例示されている意味がなくなるのではないか、と指摘されることがあります[26]。

ただ、この「ない」ことの証明問題は、実は、30条の4の構造的問題なのです[27]。これは、1号から3号に例示されていない非享受目的の場合を考えると分かります。この点、30条の4導入時から、号に例示されていない非享受目的の一例として、リバース・エンジニアリング目的があげられてきましたので、ここでもそれをもとに考えたいと思います。

リバース・エンジニアリング目的で著作物を複製する場合に、30条の4の適用を求める者は、まず、リバース・エンジニアリング目的が非享受目的であることを示さなければいけません。号に例示されていないので当然です。ただ、実はそれだけでは足りないのです。リバース・エンジニアリング目的が非享受目的であることを示した上で、その他に享受目的が「ない」ことも示さないといけません。これは使い分け説でも、単独適用説でも同じです[28]。

場合の情報解析は、非享受目的と享受目的が同時併存している場合を指すと解釈することは、30条の4第2号の「情報解析」の定義中に、「第47条の5第1項第2号において同じ」とあることに照らせば、困難であると言わざるを得ない。

[26] 愛知・前掲注13）16〜17頁および「座談会」上野＝奥邨・前掲注13）208〜209頁(愛知靖之発言)参照。

[27] 本稿でも、便宜のために、非享受目的という言葉を利用しているが、本来は正しくなく、誤解を招く表現であるのも事実である。なぜなら、30条の4は、非享受目的なら権利を制限すると定めているわけではなく、逆に、享受目的がないならば権利を制限すると定めているわけであり、「ない」ことを示す構造になっている。

[28] 単独適用説と使い分け説との違いは、各号として例示されている目的の場合に、

しかし、それは困難ですから、実際上は、リバース・エンジニアリング目的が非享受目的であることを、著作物の利用者が示せば、それ以外に享受目的があることは、権利者側が示さないといけないという具合に、責任分担させるのがバランス的に良いでしょう[29]。

仮に、各号として例示されている以外の非享受目的の場合に、上記のような責任分担が可能なのであれば、1 号～3 号の場合も、利用者の行為が、各号のいずれかに該当することが示されれば、それ以外に享受目的が併存していることは、権利者側が示す必要があるとすることが可能でしょう。そうすれば、使い分け説でも、利用者の立証上の負担が課題になることはありません[30]。

しかも、この場合も、各号例示以外の利用目的の場合は、そもそもその利用目的自体が非享受目的であることを示さないといけないのに対して、各号例示の利用目的の場合は、その必要がない点で、例示には、なお意味があることになります。

Ⅲ　30 条の 4 柱書但書について

1　議論の出発点

論点の第 2 は、30 条の 4 柱書但書(以下、単に但書と呼ぶ)に関してです。但書は、「著作権者の利益を不当に害することとなる場合はこの限りではない」と規定しています。つまり、たとえ非享受目的のみの利用であっても、但書に該当する場合は、同条の適用はなく、同条によっては権利制限されなくなるわけです[31]。

では、どういう場合が但書に該当するかですが、平成 30 年改正で現行の 30 条の 4 が導入された時点から一般的に考えられているのが、情報解

30 条の 4 が単独適用されるとするか否かだけであるため、例示されていない目的の場合、両者で差はなくなる。

[29] 権利者は、享受目的が少なくとも 1 つ併存することを示せば良い。

[30] 関連して、「座談会」上野＝奥邨・前掲注 13）219～220 頁(前田健発言)参照。

[31] 逆に言えば、享受目的の場合は、但書該当性を論じるまでもなく、30 条の 4 は適用されなくなる。

析用データベースを情報解析目的で複製する場合です[32]。これは、30条の4のもとになった旧47条の7の但書が「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」と規定していたことを踏まえて、そう考えられているわけです[33]。

「考え方について」で、但書に当たり得るとして例示されているのは、この、情報解析用データベースを情報解析目的で複製する場合に関する、3つのパターンです。逆に言うと、「考え方について」では、情報解析用データベースを情報解析目的で複製する場合以外のどういう場合が、但書に当たるかは示していません。

この関係で誤解されている場合もあるのですが、今回の「考え方について」の但書に関する議論は、従来からの議論と比べて、全く新しいことを主張したのだと理解するのは正しくなくて、従来からの解釈をより詳細にした、という位置づけだと思います[34]。

以下、3つのパターンについて見ていきましょう。

2　3つのパターン

(1) 媒体による提供もしくは電子的提供の場合

第1のパターンです[35]。まず、一般的な(つまり情報解析用に整理されているわけではない)データベースがあります。この一般的なデータベースを情報解析用に整理したもの(以下、情報解析用データベースと呼びます)が作成されます。そして、その情報解析用データベースが、DVDやUSBメモリのような媒体(メディア)に収納されて販売されたり、インターネットで電子的に販売されていたりしているとします。そのような場合に、媒体を入手したり、ダウンロードしたりして、情報解析用データベースを入手した人が、それを機械学習目的(を含む情報解析目的)で複製する場合、但書に該当する、というのが第1のパターンです。

平成30年改正直後の但書の解説は、漠然とですが、情報解析用データ

32　「基本的な考え方」前掲注3）9頁参照。
33　「基本的な考え方」前掲注3）41頁参照。
34　「考え方について」前掲注1）注27参照。
35　「考え方について」前掲注1）24頁参照。

ベースが、DVD などの媒体に収録されて販売される状況をイメージしていました[36]。一般に、データベースというのは、容量が大きくなります。平成 30 年以前の状況としては、容量の大きなものは、DVD などの媒体で提供するというのが通常でした。そういう意味では、当時の状況を踏まえると、媒体収録イメージだったのは、むしろ当然のことだったわけですが、いまや、スマートフォンでも超高速通信が可能な時代になりつつあります。結果、そこそこの容量のファイルなら、オンラインで提供することも当たり前になってしまいました。「考え方について」では、そういう状況の変化を踏まえて、DVD などの媒体による提供のケースに、オンラインでの提供を付け加えたわけです[37]。それが第 1 パターンです。

そういう意味では、そもそもの平成 30 年改正時点の解釈に反対する立場の人でもない限り、第 1 パターンについては、異論のないところだと思います。現に、「考え方について」があげた 3 つのパターンのうち、よく話題にあがるのは、第 1 パターン以外の第 2・第 3 パターンとなります。

(2) 有償のＡＰＩ経由の提供の場合

第 2 のパターンは、有償の API を経由して、データベースにアクセスする場合です[38]。ちょっと話が複雑なので、順番に考えていきましょう。なお、最初にお断りしますと、私自身は、「考え方について」を議論した審議会の小委員会に参加しておりません。そのため、今日のお話は、あくまでも、外野から「考え方について」を検討した結果となります。なかでも、この但書に関する第 2 と第 3 のパターンに関する部分は、議事録を読む限り、小委員会での議論も喧々諤々でしたし、「考え方について」の記述も分かりづらいものになっています。よって、第 2 と第 3 のパターンに関してこれからお話しする部分は、私が外野から理解する限りでの内容ということで、ご容赦ください。

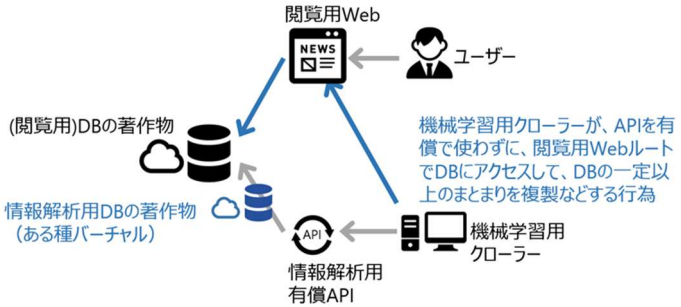
---

36 「基本的な考え方」前掲注 3）9 頁参照。

37 「考え方について」前掲注 1）24 頁参照。

38 「考え方について」前掲注 1）24～25 頁参照。

さて、話を戻します。まず、新聞社や雑誌社の web サイトを考えてください。この web サイトの実態は、記事のデータベースになります。ここで、注意喚起しておきたいのですが、新聞社などのサイトって、ばらばらに閲覧用の記事のページが存在するわけではありません。例えば、個人が、小規模な web サイトを作る際は、閲覧用のページがあって、それを樹形図的に組み合わせていく形で作成されます。そのため、webサイトの構造＝ばらばらの記事ページの樹形図的な組み合わせ、的な理解が結構浸透していたりします。しかし、新聞社とか雑誌社のサイトは何万ページもありますし、また、次々と記事が追加更新されていきます。そういうサイトを、小規模な個人サイトみたいに作成・運営することはできません。



そうではなくて、記事はデータベースに蓄積されています。私たちは、そのデータベースから引き出された記事を、フロントエンドを通じて(極端に分かりやすく喩えると、閲覧用記事ページのひな形に当てはめて)見ているんですね。そういう意味では、私たちが、判例データベースサービスを利用するときと似ています。判例はデータベースに記録されていて、その中から特定の判例を引っ張り出して、それを閲覧する。新聞社とかのサイトもそれと同じわけです。つまり、閲覧用の web サイトの実質は、記事データベースなんです。

そのデータベースから、情報解析に活用できる形でデータを引き出せる有料の API を、新聞社などが用意したとします。そうすると、ある意味バーチャルに、情報解析用のデータベースができたことになります。なぜなら、閲覧用の記事データベースに、この API を通じてアクセスするというのは、情報解析用に構成された、つまり、情報解析用の体系的

構成を持つデータベースにアクセスしていることと同義だからです。

なお、「ある意味バーチャルに」と言ったのは、閲覧用記事データベースと別の HDD に、情報解析用データベースがそれとして複製されて存在するというわけではないからです。あくまでも、HDD 上にあるのは、従来通りの閲覧用の記事データベース。それを、情報解析用の API 経由でアクセスすると、情報解析用データベースとして機能するということなのです。

そもそも、著作権法上、データベースって「論文、数値、図形その他の情報の集合物であつて、それらの情報を電子計算機を用いて検索することができるように体系的に構成したものをいう[39]」とされているだけなので、体系的に構成された形で媒体に記録されていなければならないわけではありません。体系的に情報にアクセスできるように構成されていれば良いわけで、情報解析用 API を通じたアクセスはまさにこれに当たるわけです。

ここで、注目していただきたいのは、次の 2 点です。まず、閲覧用のデータベースと、API を経由することで実現する情報解析用データベースとでは、記事の体系的構成に差があります。なぜなら、閲覧用は、閲覧に便利な体系で構成されていますが、情報解析用は、情報解析に有益な体系で構成されているからです。一方、両方のデータベースで、記事の選択は共通しています。

さて、上記のような状況のときに、機械学習用のクローラーが、インターネット上に存在する誰でも閲覧できる、新聞社の閲覧用 web サイトにアクセスして、それを一定のまとまりで複製するとします。先にお話ししたように、閲覧用の web サイトの実態は、閲覧用データベースですから、結局、閲覧用データベースを一定のまとまりで複製したことになります。

ここで、なぜ、API を経由してアクセスしないかと言うと、この API は有料で提供されているわけですから、普通は ID とか PW とかで保護されていて、アクセスできません。そのため、そういう制限がない、閲覧用の web サイト＝データベースにアクセスするわけです。

---

[39] 2 条 1 項 14 号。

先ほども申し上げましたように、閲覧用データベースと情報解析用データベースの間で、記事の選択は共通しています。と言うことは、閲覧用データベースを一定のまとまりで複製すると、閲覧用データベースの選択をコピーすることになるだけでなく、(それと共通する)情報解析用のデータベースの選択もコピーしてしまうことになるのです。

データベースの構成要素は、データ(この場合は記事)、選択、体系的構成の 3 つです。そのうち、著作物性の源泉となるのは、選択と体系的構成であり、いずれか一方に創作性があればデータベースの著作物となります。

この場合、情報解析用データベース中の、記事の選択の部分しか複製してないわけですが、それが著作物性の源泉であることは間違いないわけですので[40]、その結果、「著作物の利用市場との衝突」という視点からは、著作権者の利益を「不当に害する」と評価できる、というのが第2のパターンに関する「考え方について」の整理なのだと、私は思います。

(3) 将来販売予定の場合

最後に、最も議論の対象となっているのは、第 3 のパターンについてです[41]。

このパターンについても、例として、新聞とか雑誌の記事の閲覧用 web サイトを考えてください。先の第 2 のパターンと違うのは、現時点では、API を提供していないところです。様々な状況から、近い将来に提供する予定があることは推認されるのですが、今に関して言えば提供していない。ただ、閲覧用の web サイトには、機械学習目的などでのアクセス

---

[40] 新聞記事のデータベースの選択について、新聞に掲載された全ての記事を掲載しているはずなので、そもそも選択に創作性があるのか、疑問を持たれるかもしれない。しかしながら、編集著作物としての新聞と新聞記事データベースとの関係は、二次的著作物の関係になり得る(新聞記事データベースは、新聞における記事の選択に依拠しその同一性を維持しつつ、新たに創作的な体系的構成を付け加えたものである)から、新聞記事データベースの選択は、(原著作物である新聞由来の)創作性を有する。

[41] 「考え方について」前掲注 1) 26～27 頁参照。

を拒否する旨が、robots.txt[42]などで意思表示されているという状況です[43] [44]。

---

[42] robots.txt は技術的には一方的な意思表示に過ぎず、それに反応する（＝機械学習をしない）かどうかは、あくまでも機械学習クローラーに委ねられている。よって、著作権法上の、技術的保護手段にも技術的利用制限手段にも該当しない。この点、「考え方について」も、「技術的な措置」としか記述していない（「考え方について」前掲注 1）25 頁など参照）。

[43] 一部誤解があるので、念のために付記すると、「考え方について」は、robots.txt による意思表示があるのにそれを無視して機械学習に伴う複製を行うことが、直ちに但書に該当すると述べているわけではない。むしろ、「著作権者が反対の意思を示していることそれ自体をもって、権利制限規定の対象から除外されると解釈することは困難である。そのため、こうした意思表示があることのみをもって、法第 30 条の 4 ただし書に該当するとは考えられない。」（26 頁）と明記している。

第 3 のパターンにおいて、robots.txt の果たす役割は、情報解析用データベースが「将来販売される予定があること」を推認させる要素にすぎない。但書との関係で重要なのは、robots.txt そのものではなくて、あくまでも情報解析用データベースが将来販売予定であるという事実にある。

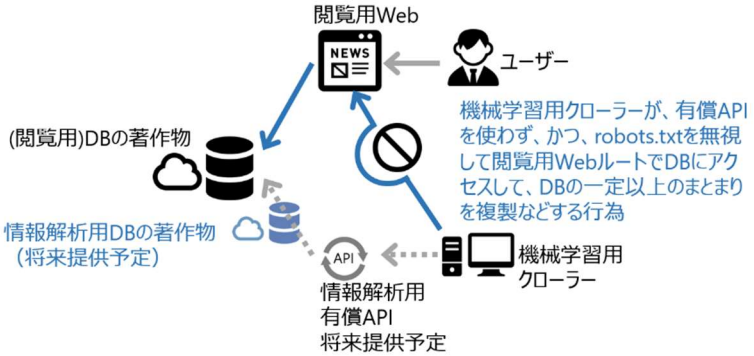
なお、「考え方について」26 頁によれば、robots.txt が前記推認の要素となるのは、robots.txt の記述が「あるウェブサイト内に掲載されている多数のデータを集積して、情報解析に活用できる形で整理したデータベースの著作物として販売する際に、当該データベースの販売市場との競合を生じさせないために講じられていると評価し得る例があ（り）」そのような場合は、robots.txt の記述は、情報解析用「データベースの販売に伴う措置、又は販売の準備行為としての措置」と言えるからである。

[44] robots.txt による意思表示があるのにそれを無視して機械学習に伴う複製を行うことが、直ちに但書に該当すると解すべきとの立場もあり得る（ただし、注 43）で述べたように、この立場は「考え方について」が採用するものではない）。

実質的にこれは、但書を通じた、robots.txt への反応強制となるが、但書の性質上、技術を特定しない一般的な強制となる。しかしながら、一方的な意思表示技術への反応を一般的な形で強制されると、強制された側の負担が過大であるばかりではなく、実現困難なことも多い。逆に言えば、反応を強制するのであれば、具体的に技術を特定する必要がある。

この点、47 条の 5 であれば、機械学習を含む情報解析の場合の、適用主体の限定要件としての政令に robots.txt を記述することで特定が可能である（現に情報検索の場合は、それが行われている）ため、筆者としては、そのような政令の追加を行い、47 条の 5 との関係では、機械学習に伴う複製を望まない権利者に拒否する（いわゆるオプトアウト）選択肢が提供できるのではないかと考えている。一方、その

こういう状況のときに、AI の開発者が、robots.txt を無視して、閲覧用の web サイトを経由し、閲覧用のデータベースを複製してしまう。そうすると、第 2 のパターンと同じで、記事の選択の部分をコピーしてしまうことになります。API をまだ提供していないので、情報解析用データベースは存在しませんが、ただ、将来登場する情報解析用データベースと閲覧用データベースで、記事の選択が共通することは分かっている。だとすると、その限りで、提供予定の情報解析用データベースの将来の販路、つまり潜在的な販路を阻害してしまうだろう、よって但書に該当するというのが、第 3 のパターンに関して「考え方について」の言いたいことだと思うのです。



(4) 小括

以上 3 つのパターンを見てきました。私自身がこれらをなんの疑問もなく納得しているかと言うと、特に第 3 のパターンについてはかなり苦しい説明だなと思っております。

また、その点を別としても、そもそも第 3 のパターンが成立するのは非常に限られた場合だろうと考えています。まず、閲覧用データベースと情報解析用データベースとで、情報の選択が共通していなければ、この説明は成立しません。つまり、閲覧用データベースを情報解析用データベースに変換する際に、なんらかの形で情報の選択に手を加えてしまい、結果、両者の間で、選択の共通性が失われてしまったとすると、第 3

ような仕組みのない 30 条の 4 では、無理である。

のパターンには当てはまらないと思います。

次に、第 3 のパターンに関して、「考え方について」では、robots.txt での意思表示を、将来の「販売の準備行為」の一環として位置づけて、情報解析用データベースが将来販売されることを推認する要素の一つにあげています。ただ、ここで言う「将来」というのは、10 年先や 20 年先というような長期的なものではないと思います。例えば、閲覧用データベースに、情報解析用 API を提供するまでの、いわば準備期間中の、権利者の救済が第 3 のパターンです。常識的に考えて、10 年も 20 年も準備期間とは言えないでしょう。また、蕎麦屋の出前のように「今から出ます、今から出ます」といつまでも引き延ばすことも許されないでしょう。実際に、過去の事例に照らして、情報解析用のデータベースを提供するまでに、どれぐらいの期間が必要だったかなどを踏まえて、どの程度の準備期間が必要かが判断されるものと思います。具体的な進展がなく、いつまでも情報解析用データベースが提供されない場合には、この推認は成立しないと考えています。

(5) あり得る誤解について

但書についてですが、ネットなどを見ていますと、一部に誤解されて理解されているところがあると思います。しかも、文化庁から公開された、「考え方について」を分かりやすく説明した「AI と著作権に関するチェックリスト&ガイダンス[45]」（以下、チェックリスト&ガイダンス）中に、それを加速するような記述があって、気になっています。

「チェックリスト&ガイダンス」の 40 頁では、まず、「『AI 学習等の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』（AI 学習用データベース）が販売されている場合、『この AI 学習用データベースを、AI 学習等の情報解析目的で複製等すること』は・・・『著作権者の利益を不当に害することとなる場合』に当たる」とした上で、「権利者としては、インターネット上に自らの作品を公開する場合、その作品を含んだ AI 学習用データベース（データセット）を、AI 学習等の情報解析目的で販売

[45] 文化庁著作権課「AI と著作権に関するチェックリスト&ガイダンス　令和 6 年 7 月 31 日」。

（ライセンス提供）するようにしておくことで、無許諾で学習されることを法的に防ぐことが可能です。」（下線筆者）との記述があります。

前半は良いのですが、問題は後半です。と言うのも、下線を引いた部分、実は言葉が足りないんです。すなわち、「何が」無許諾で学習されることを防ぐことができるのかが記されていないのです。データベースの複製が禁止されるのか、それとも著作物そのものの複製が禁止されるのか、どちらなのかが曖昧なのです。

そのため、ネット上では「あるサイトで自分の著作物Aを AI 学習用データベースとして販売しておけば、他のサイトに掲載した著作物Aも機械学習によって複製されない」と思っている方が結構いるように思われます。

しかし、これは誤解です。先ほどからご説明したように、「考え方について」が述べているのは、あくまでも、情報解析用データベースは、データベースとしては機械学習目的で複製されません[46]、ということを述べているだけです。情報解析用データベースを構成する素材である著作物が、他のサイトに―情報解析用データベースとしてではなく―掲載されている場合は、それ自体を、機械学習目的で複製することには、元通り 30 条の 4 の適用があるのです。

個人的には、誤解が広まるのは、良くないと思っています。誤解して、「学習されない」と安心していたら実際には学習されてしまう、という状況になりかねないからです。

「チェックリスト＆ガイダンス」のもとになった「考え方について」の該当する部分を見ると、そちらは誤解が生じないように記述されています。「チェックリスト＆ガイダンス」として、表現を簡略化する際に、「何が」に当たる言葉を落としてしまったのが問題だと思います。

(6) 果たして但書の問題か

ここまで、情報解析用データベースを情報解析目的で複製することは、但書に該当するということを前提に、「考え方について」が示した 3 つの

[46] 正確に言うと、情報解析用データベースを、データベースとして機械学習目的で複製すると、但書に該当して 30 条の 4 の適用がなくなるため、他の権利制限規定の適用がない限り、無許諾で行うと著作権侵害となるということ。

パターンに関してお話ししてきました。ただ、そもそも論になってしまいますが、私自身は、本当に但書の問題なのだろうかという疑問を感じています。

この点については、神戸大学の前田健先生もおっしゃっておられて、私も同じ意見なのですが、情報解析用データベースを情報解析目的で利用するというのは、そもそも、享受目的だろうと思うのです[47]。

享受とは、端的に言えば、著作物の本来的な利用のことです。例えば、絵画であれば「見ること」、小説なら「読むこと」、コンピュータ・プログラムは「実行すること」が享受に当たります。では、情報解析用データベースの場合の享受とは何でしょうか？ 情報解析用データベースの本来的な利用は、「情報解析に使うこと」ですから、それこそが享受に当たるわけです。

としますと、第 1 のパターン、すなわち、市販されている情報解析用データベースの DVD や、ダウンロード販売されている情報解析用データベースを、情報解析目的でコピーする行為は、明らかに享受目的の行為ですから、但書が登場する余地はなく[48]、30 条の 4 は適用されません。

第 2・第 3 のパターンの場合は、情報解析用データベースの特徴である、情報解析用の体系的な構成を利用していないわけですから、情報解析用データベースをその本来の目的に沿って利用している、つまり享受していることにはなりません。つまり、非享受目的の利用だと思います。したがって、これらの行為は、本来は、30 条の 4 の適用対象になり得るのではないかと、私は思っております。もちろん、これは「考え方について」とは異なる見方です。

(7) オルタナティブ

「考え方について」の第 3 のパターンの本来的な狙いは、robots.txt によるオプトアウトを、30 条の 4 の枠内で、多少なりとも実現しようとする

---

[47] 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(2)著作権法編』（勁草書房・2023）207～208 頁参照。

[48] 但書該当性が問題になるのは、非享受目的の場合であり、享受目的の場合はそもそも問題になりようがない。

ところにあるのだと思います。そうすると、30 条の 4 の条文上手がかりになりそうなのは、但書しかないため、そこでなんとか読み込もうと苦労しているわけです。しかし、先ほども申し上げましたように、それはかなり苦しい説明だなと思います。

ところで、機械学習を含む情報解析について規定した条文としては、30 条の 4 だけではなくて、47 条の 5 も存在することは、先ほどご説明したところですが、実は、47 条の 5 ならば、robots.txt で学習を禁止することが可能です[49]。

47 条の 5 は「情報検索」と「情報解析」のための規定ですが、現在、「情報検索」に関しては、robots.txt を遵守することが、47 条の 5 の適用条件として政令で定められています[50]。一方で、「情報解析」には、相当する政令がありません。なぜかと言うと、47 条の 5 導入時点では、情報解析に関して、オプトアウトを可能とするような標準的な技術的手段がなかったからです。

しかし、47 条の 5 が導入された後、社会は大きく変わりました。インターネットから情報を収集して、機械学習の対象とすることが一般的になりました。また、現時点では、robots.txt は、検索エンジンだけでなく、AI の学習についても、オプトアウトする手段として一般化しつつあります[51]。そうであるならば、情報解析に関しても政令を整備し、robots.txt を遵守しなければ、情報解析＝機械学習について 47 条の 5 の適用が受けられないとしても良いと思うのです。

しかも、47 条の 5 第 1 項の但書は、①権利者の利益を不当に害する場

---

[49] 正確には、robots.txt で、学習に伴う複製などを 47 条の 5 の適用対象外とすることが可能。

[50] 付随的軽微利用（法 47 条の 5 第 1 項）の主体については、著作権法施行令 7 条の 4 第 1 項。準備行為のための複製（法同条 2 項）の主体については、施行令 7 条の 4 第 2 項。

[51] 例えば、OpenAI 社は、同社の AI 用の web サイトクローラーを robots.txt でブロックする方法を公開している（宇都宮充「OpenAI、AI 学習用 Web クローラーのブロック方法を紹介」PC Watch, 2023 年 8 月 8 日 17：57, https://pc.watch.impress.co.jp/docs/news/1522758.html）。他の大手 AI 開発会社も同様の取組みをしていると報道されている。

合、の他に、②海賊版であると知りながら付随的軽微利用する場合も規定しています。つまり、先の政令整備によるオプトアウト対応とあわせると、EU の AI 法（AI Act）と似たことが、47 条の 5 を活用することで実現可能なわけです。30 条の 4 ではできない芸当であり、単独適用説ではそもそも不可能です。

ただ、そのためには 2 つ課題があります。

まず、47 条の 5 は享受目的が併存する場合の規定なので、47 条の 5 を活用しようとすると、享受目的の併存を広く認定しなければなりません。享受目的の併存を広く認めると、30 条の 4 の適用範囲が狭まり、47 条の 5 に誘導されることになります。

次に、47 条の 5 で許される享受目的利用は、「付随的」軽微利用でなければならないという点です。例えば、先ほどのピカソ作とされる絵画の真贋判定 AI の例のように、単純なパーセント表示だけでも真贋判定は成立するのだけれど、サムネイル画像があった方が分かりやすいというような場合は、付随的だと言えます。47 条の 5 立法直後から、文化庁も似たような説明しておりますし、私もこれまでそういう立場をとってきました。しかし、この考え方でいくと、生成 AI による出力がこの要件を満たすケースは非常に少なくなってしまいます。

文化庁や私のような考え方は「外側説」と呼ばれていますが[52]、これに対して、弁護士の松田政行先生などがとられている「内側説」という立場があります[53]。今日は時間がないので、内側説の中身については割愛しますが、内側説を採用すると、生成 AI の出力で、付随的と言える範囲が拡がるかもしれません。そうだとすると、「外側説」から「内側説」へ、私自身は宗旨替えをしても良いかなと思っています。

私はあまり自説に固執しないタイプですので、以前の見解を「間違っていました、ごめんなさい」と言って変えることがあります。今回も、そろそろ「ごめんなさい」と言って内側説に変更するかもしれません。と言うのも、そうすることで、47 条の 5 の適用範囲を広げ、結果的に

---

[52] 「基本的な考え方」前掲注 3）22 頁参照。

[53] 「外側説」「内側説」については、加戸守行ほか「〔座談会〕平成 30 年改正著作権法施行に伴う柔軟な権利制限規定による著作物の利用拡大とこれからの課題（下）」NBL1145 号（2019）33～35 頁参照。

robots.txt や海賊版対策も、無理をしなくても対応できるようになるのではないか、と個人的には思っているからです。もちろん、まだ未検証の部分が多いので、今後も慎重に検討していく必要はありますが、現在考えているのはそういった方向性です。

Ⅳ　依拠について

1　2 段の依拠

3 つの論点の最後は、依拠についてです。

私は、生成 AI において依拠を考える際には、2 段階の検討が必要だと思っています[54]。「考え方について」もこの点は一致です。

なお、私はこの 2 段階の依拠を、民訴における 2 段の推定に擬えて、2 段の依拠と言っています。具体的には、AI 操作者による依拠と AI による依拠です。表にもあるように、AI 生成表現が、既存の著作物（の創作的表現）に類似している場合、当該既存の著作物に、操作者が依拠している場合（操作者による依拠）も、AI が依拠している場合（AI による依拠）も、どちらでも侵害が成立します。逆に言うと、侵害が成立しないのは、両方の依拠が存在しない場合だけです。

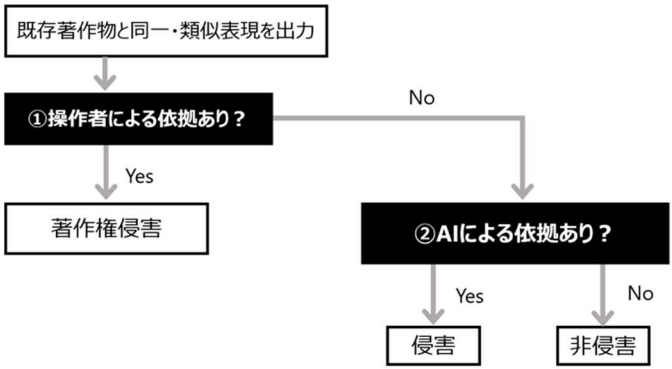
AI生成表現が既存の著作物に類似する場合

		操作者による依拠	
		あり	なし
AIによる依拠	あり	侵害	侵害
	なし	侵害	非侵害

[54] 本論点についての詳細は、奥邨弘司「依拠・類似」上野＝奥邨・前掲注 13）108～127 頁参照。

２　操作者による依拠

まず、2 段の依拠の検討順序ですが、理論的にはどちらが先でもかまいません。先ほど申し上げたように、操作者による依拠を先に検討して、次に AI による依拠を検討するのでも、逆に、AI による依拠を先に検討して、その後に操作者による依拠を検討するのでも、どちらでも良いのです。重要なのは、順番は別として、両方を検討する必要があるという点です。



ただ、現在の生成 AI を取り巻く状況、特に、生成された表現が既存の著作物に類似しているという問題に対処する上では、操作者による依拠を先に検討した方が、効率が良いのは事実です。と言うのも、現状では操作者による依拠を証明する方が簡単です。なぜなら、これまで、つまり生成 AI が登場するまでと同じように考えればよいからです。

例えば、任天堂のマリオを過去に見たことがある操作者が、「任天堂マリオ」と画像生成 AI に入力して、そっくりな画像が出力されるまで、何回も何回も繰り返し入力と出力を繰り返したような場合は、過去に見たことがある画像と似たものを意図的に作っているわけですから、依拠があると言えます。

また、過去にスタートレックのエンタープライズ号の映像を見たことのある操作者が、「宇宙船」と画像生成 AI に入力して、たくさん出力されたうちから、エンタープライズ号に似たデザインのものを意図的に選

んでネットに公開したような場合はどうでしょう。生成自体は、AI が偶然に行ったかもしれませんが、エンタープライズ号に似ていることを理解した上で、それを敢えてインターネットで公開したならば、少なくとも公衆送信権侵害の文脈では、依拠ありです[55]。

結局、AI の代わりに、鉛筆や、ペン、パソコン、Photoshop などの道具を使った場合と同様に考えればよいわけです。そういう意味で、操作者

---

[55] ところで、人間の場合、創作された表現が、既存の著作物と(創作的表現において。以下略)類似していたとしても、その創作者に関して独自創作が成り立つ(＝当該創作者は、類似する既存の著作物の存在も内容も知らなかった)場合、その表現を頒布したり公衆送信したりする第三者が、類似する既存の著作物を知っていたとしても、当該第三者について著作権侵害は成立しないとされる(田村善之「著作権法概説　第2版」(有斐閣・2001) 55〜56 頁参照)。

これとのアナロジーで、学習対象著作物に含まれない既存著作物と類似する表現を AI が生成した場合に関して、操作者が意図的に出力させたのではないなら、AI による独自創作と解し、操作者が類似する既存の著作物を知っていても、その頒布や公衆送信について、著作権侵害は成立しないと解すべきではないかとする見解がある。操作者を、前記人間の場合の例の第三者に相当すると考えるわけである。

しかしながら、この場合も、操作者がAIを用いて表現を作成している以上(ここでは、AI 生成表現の著作物性を論じているわけではないため、操作者による創作的寄与が認められるような態様での操作の場合は当然、それが認められないような態様の操作の場合も含む)、独自創作の成否は、操作者を基準に判断すべきである。もっとも、厳密に言えば、操作者が意図せず、(操作者が知っている)既存の著作物と類似する表現が AI から出力された場合、まさにその瞬間に関して、操作者に著作権侵害責任を問うのは酷であるが、既存の著作物と類似することを知りつつ頒布や公衆送信を行った場合は、意図的に生成したのと同様に捉えて問題ないであろう。

以上のように解さないと、既存の著作物と類似する表現を、AI を操作して出力し、頒布や公衆送信している者が、当該既存の著作物を知っていても、意図して AI に出力させたわけではないと反論する余地を残してしまい、それこそ著作権侵害ロンダリングを生みかねない。

なお、筆者の立場でも、AI を操作した者と、それを頒布・公衆送信する者が別の場合、前者について独自創作が成立すれば、後者が類似する既存の著作物を知っていても、頒布・公衆送信について著作権侵害は成立しないと解する。

本注に関して、上野＝奥邨・前掲注 13) 276〜285 頁の議論も参照。

による依拠というのは、今までと変わらないので、AI の文脈特有の複雑なことはありません。一方で、後で詳しく見ますが、AI による依拠を検討するためには、AI が、何をどのように学習したのかを調べる必要がありますが、それは非常に困難で複雑です。

現在、AI が出力した表現について、既存の著作物の著作権を侵害しているのではないかと、インターネットなどで問題となっている事案は、ほとんどが、操作者による依拠が認められるケースです。だとすると、操作者による依拠をまず吟味すれば、AI による依拠というややこしい問題に取り組む必要がないわけです[56]。

今、インターネットのことを話題に出しましたので、ついでに申し上げますと、ネット上では時々、被疑侵害表現を作成した人が、「これは AI が作成した画像なので著作権侵害ではありません」みたいな主張をしていることがあって、それに対して「著作権(侵害)ロンダリング」なんて呼ばれることもあります。しかし、既存の著作物と類似の表現を、AI を用いて意図的に作成したなら、操作者による依拠が認められますから、普通に著作権侵害であり、全然ロンダリングなんかできていません。

３　ＡＩによる依拠

(1) 基本型

AI を用いようが、用いまいが、表現の作出者が、過去に見たことや聞いたことがある著作物に、意図的に似せて表現を作成すれば、依拠があります。その場合、AI を用いていたのなら、ここで言う操作者による依拠が認められます。

では、操作者が、見たことも聞いたこともない既存の著作物と類似する表現が AI から出力された場合はどうでしょうか。この場合は、ワン・レイニー・ナイト・イン・トーキョー事件最高裁判決[57]に照らせば、操作者による依拠は認められません。それは、そのようにして出力されたものを、インターネットで公衆送信した場合も同様です。

---

[56] もっとも、AI 開発者の責任を議論する場合は、AI による依拠＝学習の有無を明らかにする必要がある。

[57] 最判昭和 53 年 9 月 7 日民集 32 巻 6 号 1145 頁。

しかしながら、これで依拠の検討が終わるわけではないのです。まだ1段目の操作者による依拠の検討が終わっただけであり、それが認められなかった場合は、2段目のAIによる依拠を検討する必要があります。

AIによる依拠の判断基準の基本型は、割と簡単明瞭で、AIが学習していなかった場合は「依拠なし」、学習していた場合は「依拠あり」となります。これが基本型です[58]。ただ、「考え方について」は、この基本型に加えて、「学習に用いられた著作物の創作的表現が、生成・利用段階において出力される状態となっていないと法的に評価できる場合」は依拠が認められないと述べています[59]。つまり、先の基本型の例外の場合があると述べています。

「学習に用いられた著作物の創作的表現が、生成・利用段階において出力される状態となっていないと法的に評価できる場合」が、何を意味するかは後ほど見ていきますが、その前に申し上げておきますと、私は理論的にそういう場合が確実にあると考えておりますが、一方で、実務的には、なかなかそれを証明するのは難しいのではないかとも思っています。

と言うのも、AIによる依拠＝AIが学習したかどうかを、権利者が立証するのは通常困難です。よって、実務的には、間接的な事実から推認される形になると思います。これは、従来、依拠を証明してきた場合と同じです。結果、例えば、類似度が高い表現が高頻度または少ない施行回数で出力される場合、通常は、AIによる依拠がある＝AIが学習していたと推認されることになると思います。

もちろん、推認ですから、AIの開発者や操作者が、AIの仕組みを一生懸命説明して、「学習に用いられた著作物の創作的表現が、生成・利用段階において出力される状態となっていない」と具体的に証明できるならば、推認を覆す余地を残しておくべきというのが、「考え方について」の趣旨なのではないかなと考えていますし、私はそれに賛成です。

とは言え、実際には間接事実による推認を覆せないまま、訴訟が終わるケースがほとんどではないかとも思っています。したがって、ここか

58　「考え方について」前掲注1)34頁参照。

59　同上。

ら先はかなり理論的な話と、ご理解ください。

(2) 例外

では、「考え方について」が言う、「学習に用いられた著作物の創作的表現が、生成・利用段階において出力される状態となっていないと法的に評価できる場合」とはどういう場合でしょうか。

「考え方について」は「開発・学習段階において学習に用いられた著作物の創作的表現が、生成・利用段階において生成されることはないといえるような状態が技術的に担保されている」などの事情が認められる場合、前記のような法的評価ができるとしています。そして、技術的に担保されている場合の具体例として、①「学習に用いられた著作物と創作的表現が共通した生成物が出力されないよう出力段階においてフィルタリングを行う措置が取られている場合」と②「当該生成 AI の全体の仕組み等に基づき、学習に用いられた著作物の創作的表現が生成・利用段階において生成されないことが合理的に説明可能な場合」をあげています[60]。

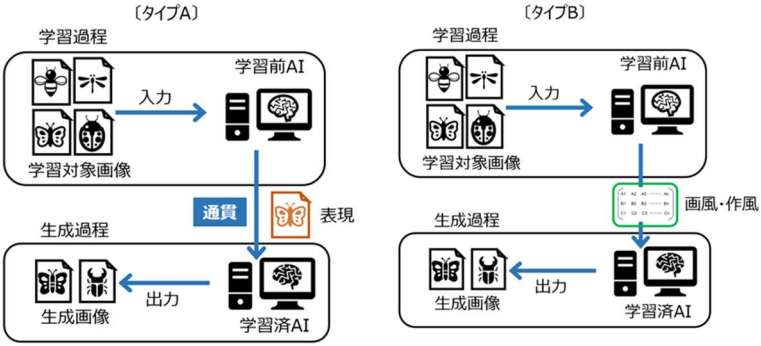
①のフィルタリングの場合は、分かりやすいと思います。問題は、②の全体の仕組みから合理的に説明可能な場合です。これは一体、何を意味するのでしょうか。

ここで、改めて、「考え方について」が述べているところをご覧いただきたいのですが、「考え方について」は、「出力されない場合」ではなくて、「出力される状態となっていないと法的に評価できる場合」と記述しています。どういうことかと言いますと、100%出力されないわけではない、出力されることも、ごくまれにはあるのだけれども、それでも、法的には「出力される状態となっていない」と評価できる場合を問題にしているということです。

ここからは、田村先生と私とで、考え方が異なる部分です。

---

[60] 「考え方について」前掲注 1)注 45。



私は、AIには大きく分けて2つの種類があると考えています。1つは、学習対象著作物の創作的な表現（絵柄や文体）を理解し、それをモデル内になんらかの形で保存して、再現的に出力するタイプ（以下、タイプA）です。この場合、生成された表現中に学習対象著作物の創作的表現と類似する部分が存在すると、学習対象著作物の創作的表現が生成過程でも利用されていると言えます。つまり、学習対象著作物が、学習過程と創作過程の間の一種のバリアを越えて、学習過程だけではなくて、生成過程でも使われてしまっているという状況です。これを私は「通貫現象[61]」と呼んでいます。学習過程と生成過程の間のバリアを貫いているので「通貫」です。

一方で、学習対象著作物の表現（絵柄や文体）からアイデア（画風や作風）を抽出して、アイデアだけをモデル内に保存しているタイプの場合、当然ですが、学習対象著作物の創作的な表現は、モデル内には保存されていません。このタイプ（タイプB）の AI の場合に、出力された表現内に、仮に、学習対象著作物の創作的表現に類似する部分が認められたとしても、それはアイデアである画風や作風が似ていることに起因する、偶然の結果だと言えるでしょう。

そもそも、出力表現が、学習対象著作物とアイデア（画風・作風）のレベルで似ていても、創作的表現まで似ることは、通常ならばありません。

[61] もしかすると、日本語として「貫通」の方が自然かもしれないが、「貫通」という言葉は法律の分野では（漢字と意味は違うが音の）似た表現もあるため、誤解を招かないように、ここでは「通貫」と表現している。

しかしながら、画風や作風が似ていると、例えば 1 万回に一度などの確率で、本当に偶然に、創作的表現である絵柄や文体まで似てしまうことがあり得ます。画風や作風が似ているわけですから、多数回繰り返せば、そういう可能性は否定できないでしょう。そして、もしこれが「本当に偶然」であることが証明できれば、非侵害とすべきでしょう。なぜなら、通貫現象は発生していないからです。学習対象著作物の創作的表現が、モデル内に記憶されていない以上、学習過程と生成過程の間のバリアは維持されています。

整理しましょう。

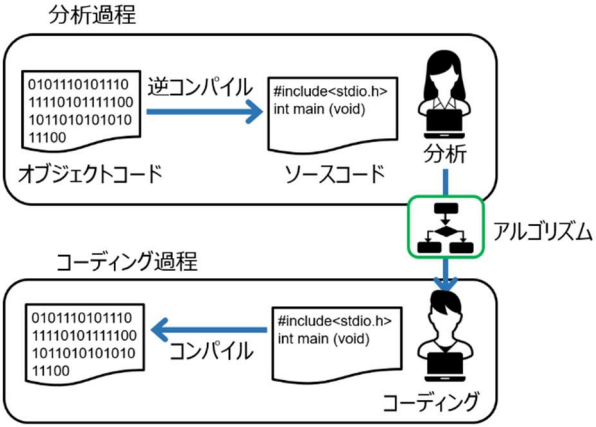
生成 AI から、既存の著作物と創作的表現が類似する表現が出力された場合、通常は、その既存の著作物を AI が学習していた＝依拠があったと推認されます。この場合、AI のタイプは、タイプＡであることが前提になっています。しかし、もし AI の開発者や利用者が、技術的な仕組みなどに基づいて、問題の AI が、タイプＢであることを立証できたとすると、類似する表現が出力されたのは、真に偶然と言うことになり、依拠は否定されます。「考え方について」が言う、「学習に用いられた著作物の創作的表現が、生成・利用段階において出力される状態となっていないと法的に評価できる場合」とは、まさにこういう状況だと私は思うのです。

(3) リバース・エンジニアリングとの相似

タイプＡとタイプＢに分ける考え方、特にタイプＢの場合に依拠が成立しないという考え方は、実は、コンピュータ・プログラムのリバース・エンジニアリングにおける「クリーンルーム方式」に関する議論と、基本的な考え方は同じです。

クリーンルーム方式のリバース・エンジニアリングでは、プログラムの開発チームを、既存のソースコードの分析だけを担当する分析班と、新しいプログラムを作るコーディング班に分けます。既存のソースコードを見るのは、分析班だけとし、コーディング班は既存のソースコードを直接見ることはしません。分析班とコーディング班の間の、情報のやりとりは厳格に管理し(チャイナ・ウォールを立てるとも言います)、コーディング班は、分析班が既存のソースコードから抽出したアルゴリズム＝アイデアだけを受け取って、それをもとに新しいプログラムを作成

します。この場合、仮に、コーディング班が、既存のソースコードと創作的表現が類似するソースコードを作成してしまったとしても、コーディング班は、アイデアに依拠しただけで創作的表現への依拠はないので、著作権侵害には当たらないとされます。



私は、これと同様のことが、生成 AI の場合にも言えるのではないかと考えています。具体的には、クリーンルーム方式の場合の分析班が、生成 AI の場合は学習過程に当たり、コーディング班が、生成過程に当たります。それで、タイプＢの AI の場合は、クリーンルーム方式そのままで、チャイナ・ウォールが完全に機能しているために、学習過程（＝分析班）から生成過程（＝コーディング班）に引き渡されるのは、学習対象著作物のアイデアのみとなり、生成過程はそのアイデアだけに基づいて表現を生成するので依拠なしとなります。

一方、タイプＡの場合は、チャイナ・ウォールが十分に機能してない状態です。そのため、学習過程から生成過程に、学習対象著作物のアイデアだけでなくて、創作的表現も（一部）引き渡されてしまう、つまり通貫が生じてしまう。その結果、生成過程は、学習対象著作物の創作的表現にも基づいて、表現を生成することになりますので、依拠ありという結論になるわけです。

(4) 批判

先ほど、田村先生と私で考え方が違うと申し上げたポイントは、実は、このクリーンルーム方式による場合に、依拠がないとする結論を肯定するか否定するかという点にあります。ここまでお話ししたように、私は肯定するわけですが、田村先生は否定されます。その結果として、田村先生のお立場だと、タイプＢのような AI についても依拠は否定されなくなります。

私が理解する限り、田村説では、依拠要件は、独自創作者が不測の損害を被ることを防ぐところにあるところ[62]、クリーンルーム方式を採用している場合でも、既存のソースコードを知った上で新しいコードを作成する場合は、不測の損害が生じるとは言い難く、結果、依拠要件で保護する必要はない、ということだったかと思います[63]。さらに、1 つの会社内部で、分析班とコーディング班に分けても、会社＝法人単位で考えると、既存のソースコードに依拠していると言えるのではないか、というご指摘もあったように思います[64]。

もっとも、田村説でも、クリーンルーム方式に全く意味がないというわけではなくて、クリーンルーム方式は、依拠との関係ではなくて類似性との関係で意味があるとされます。つまり、クリーンルーム方式を採用すれば、普通は似たコードが作成されることはないし、仮に似たコードが作成されるとしても、それは、誰が書いても似たコードになるということを意味していて、類似性否定の心証が形成されるだろう、と説かれます[65]。

(5) 批判への疑問

以上に対しては、なるほどと思う一方で、いくつか疑問も持っています[66]。

---

62 田村・前掲注 55)48～49 頁参照。

63 田村・前掲注 55)55～56 頁参照。

64 田村・前掲注 55)56 頁参照。

65 同上。

66 本文にあげた以外の疑問点として、分析班とコーディング班を会社内で分けても、会社＝法人単位で見れば、依拠ありと言えるとの指摘については、例えば、

まずは、実際的な問題です。クリーンルーム方式によって、依拠が否定されなくても、類似性が否定されるのであれば、結局は著作権侵害にならないので、問題がないのではと思われる方もいるかもしれません。

しかし、依拠で否定するのは、入り口で否定することになりかなり明確です。具体的には、クリーンルーム方式をきっちり運営しているということさえ示せれば、既存のソースコードと作成されたコードが似ているか似ていないかに踏み込むことなく、侵害が否定されます。実務的に言うと、クリーンルーム方式の体制をきっちり運営し、そのことを信頼に足る記録として残しておけば、安心してリバース・エンジニアリングができるわけです。できあがったソースコードを吟味したりする必要はありません。

一方で、類似性の否定の場合は、そうはいきません。新しく作成されたソースコード中に、既存のソースコードと似ているものがある場合は、それが本当に、誰が書いても似るものかが問題となり得ます。確かに、クリーンルーム方式で作成したコードだということを示せば、似ていたとしても侵害にはならないということを示す 1 つの要素にはなるでしょう。しかし、例えば訴訟の場面を考えた場合、既存のソースコードの権利者が、それで納得するかというと、普通は、クリーンルーム方式で作成しても、別のコードが作成される余地はあると示してくるでしょう。となると、リバースした側としては、反論し、さらに再反論、再再反論という具合に繰り返しです。しかもそれを、似ているところ全部についてしなければならない。

となると、実務的には、予め、できあがったソースコードを、既存のコードと比べて、似ているところをチェックし、誰が書いても似てしまうと、100%の自信で言い切れるところ以外は、似ていないように書き直さないといけないということになりかねません。これは実務としては、大きな負担で、正直リバース・エンジニアリングを行うのは困難になってしまいます。

次に、理論的な点です。

私は、依拠の要件を次のように位置づけています。そもそも、著作権

分析班を外注に出せばクリアすることが可能ではないかと考える。

は、特定の表現について独占権を与え、国家機能である司法を通じてそれに実効性を与えるものですから、どう贔屓目に捉えても、（著作権者から見て）他人の表現行為を規制する性質を有しています。この著作権の持つ本質的な危険性を踏まえると、単に既存の表現と同一または類似であるというだけで権利侵害になるとすると、著作権者以外の表現活動を大きく萎縮させかねません。そこで、侵害の成立範囲は、著作権者の表現の無断利用の場合に限られるのだと思います。著作権が相対的独占権であるのはまさにこういう理由です。

ところで、著作権者の表現の無断利用であることが、明らかな場合もあれば、分かりづらい場合もあります。依拠要件は、著作権者の表現の無断利用であることを示すのが難しい場合にこそ意味を持つものなのだと考えます。逆に言うと、著作権者の表現の無断利用であることが明らかな場合、敢えて依拠があったかどうかを分析する必要も、言及する必要もないのです。

例えば、①コピー機でコピーされた絵画が、著作権者の絵画と同一の場合や、②CD を再生した音楽が、著作権者の音楽と同一の場合を考えてください。いずれの場合も、機械の仕組み上、入力されたものが出力されますので、結果物（出力されたコピーや再生された音楽）が著作権者の著作物と同一であれば、著作権者の著作物が利用されたと結論できます。よって、このときは敢えて、依拠を論じるまでもありません。

一方で、③画家が描いた絵が、著作権者の絵画と同一の場合や、④演奏家が楽器で演奏した音楽が、著作権者の音楽と同一の場合は、どうでしょうか。結果物（描かれた絵や演奏された音楽）からは、例えば、画家が、著作権者の絵画を（現に見ながら、または過去に見たものを思い出しながら）模写したのか、それとも自分で閃いて描いたのか（＝独自創作したのか）、それを区別することはできません。こういう場合に、依拠の議論が出てくるわけです。

議論を AI の場合に戻しましょう。

生成 AI の場合、①や②のコピー機や CD 再生機の場合のように、仕組み上、入力されたものが出力されると言い切ることができるでしょうか。この点、生成された表現と類似している著作物が学習されていれば、常に依拠ありとする立場は、生成 AI をコピー機などと同様に考えているこ

とになります。そういう意味では、この立場の場合は、依拠について敢えて言及する必要すらないようにも思います。

一方で、私は、生成 AI をタイプＡとタイプＢに分ける立場です。私の立場でも、タイプＡの場合はコピー機などと同じと考えます。なぜなら、学習過程から生成過程への通貫があり、入力されたものが出力されているからです。ただ、タイプＢの場合は、学習過程から生成過程への通貫がありませんから、入力されたものが出力されているとは言えず、コピー機などとは異なると考えることになります。その上で、タイプＢの場合は、人間で言うと、画家が自ら閃いて描いた場合（＝独自創作）と同じように考えるべき、つまり、依拠はないと考えるべきだと思っているわけです。

ここで多分、タイプＢとコピー機が異なるというのは同意するが、タイプＢを独自創作と評価することには違和感があるという方が多いのではないでしょうか。なぜなら、人間の場合は、過去に見聞きしたものと同一の表現を作成した場合に、自分で閃いた＝独自創作したという主張は容れられず、依拠があるとされるのが普通だからです。

ただ、人間の場合に、今説明したような状況で、独自創作の主張が容れられないというのは、実は、理論的な話ではないと思っています。と言うのも、人間の場合の学習にも、本来は、タイプＡとタイプＢがあって、過去に見聞きした著作物をタイプＢで学習していたのなら、理論的には依拠が否定されるべきだと思うからです。しかしながら、人間の脳内は、今の科学では未解明です。そのため、タイプＡとタイプＢを区別することができません。そして、多くの人の実感として、タイプＡ的な学習が多いことを知っています。そのため、ある意味経験則に基づき、人間の場合は、タイプＡの学習なのだとして処理されているにすぎないと思うわけです。

しかし、AI の場合は、タイプＡとタイプＢを技術的に区別できる可能性があります。もし、本当に、両者を区別できるのならば、タイプＢの場合は、アイデアに基づいているだけですから、依拠は否定されるべきなのです[67]。

[67] その意味では、AI の場合だけではなくて、人間の場合も、例えば、脳の病気で、

もっとも、誤解を避けるために申し上げると、AI の場合も、今はその仕組みにブラックボックスのところが多いと言われています。ですので、タイプＡとタイプＢの区別は難しく、また、現状は、学習対象著作物の創作的表現そのものを、なんらかの形で保存していないと、それと同一のものが出力される可能性は極めて低いということに照らすと、タイプＢであることが明らかに示されない限り、タイプＡとして扱うのはやむを得ないと思います。また、タイプＢと示せる場合は、人間の場合よりは多いかもしれませんが、それでも、非常に限られた場合でしかないとも思っています。

ただ、もしタイプＢと示せるのなら、そういう場合は依拠を否定する理論的な道筋を残しておくべきだとも思っていて、だからこそ先ほどから縷々お話ししているような次第となります。今後、AI の技術がどう進歩するかも分からない中で、現状だけを前提に、AI が学習している以上常に依拠はある、と結論づけてしまうのは、やっぱり早計で、色々と弊害があるだろうと思っております。

以　上[68]

［付記］

本稿には、JSPS 科研費 JP24H00131 の助成を受けた研究の成果が含まれている。

過去に見たり聞いたりしたことを完全に忘却してしまう状態にあることが医学的に証明できたなら、その場合に創作された過去に見たり聞いたりしたものと同一・類似の表現については、独自創作として扱うべきだろう。

[68] 図中のアイコンは、Icon-rainbow（https://icon-rainbow.com/）および ICOON MONO（https://icooon-mono.com/）のものを利用した。記して感謝したい。