

人工知能による画像生成物の著作物性と表現脳

鈴木 仁志*

I. 序論

近時急激な進化を遂げている生成 AI (Generative AI) に関しては、著作権との関係で様々な議論がなされているが、中でも、生成 AI の利用者 (以下、単に「利用者」という。) が AI に指示を入力した結果として出力された生成物は利用者の著作物となるか (著作物性、利用者の著作者性) との点 (以下「本論点」という。) については、AI に対応した特別立法が行われていないことから、既存の著作権法の枠組みの中で様々な解釈論が主張され¹⁾、文化庁の文化審議会著作権分科会においても議論が行われている。

この点、AI は、脳の神経回路を模したニューラルネットワーク (Neural Network) を用いる「人工知能」であるから、利用者の創作的表現性の有無について判断するためには、創作的表現を行う人間の脳の機能と生成 AI の「知能」としての機能との関係を具体的に検討することが有用と考えられるが、筆者の知る限り、このような試みはこれまであまりなされてこなかったように思われる。

そこで、本稿では、社会的に先鋭的な議論の見られる描画の問題²⁾ を題材として、人間が描画する場合と利用者が言語指令によって AI に画像を生成させる場合とを対比しながら、創作的表現を行う脳機能と画像生成 AI の仕組み及びそれらの関係性を分析することにより、画像生成物の著作物性についての考察を行う。

II. 問題提起 — 「道具」の概念と人工知能

後述の通り、文化庁は、本論点について、「人

が思想感情を創作的に表現するための『道具』として AI を使用したものと認められれば、著作物に該当し、AI 利用者が著作者となる」とし、「人が AI を『道具』として使用したといえるか否かは、人の『創作意図』があるか、及び、人が『創作的寄与』と認められる行為を行ったか、によって判断され (る)」と解した上³⁾、「創作的寄与」の該当性 (生成 AI に対する指示の具体性と AI 生成物の著作物性との関係) について、①指示・入力 (プロンプト等) の分量・内容、②生成の試行回数、③複数の生成物からの選択等の諸要素を考慮して判断するものとしている⁴⁾。

しかし、AI は機械であり、「物を作り、また事を行うのに用いる器具」であるから、人間にとっていかなる意味においても「道具⁵⁾」である。したがって、AI を「道具」として使用した場合には利用者が著作権を享受するとの一般的な判断基準には検証の余地がある。

この点、著作権法上の他の議論における「道具」の概念は、他人をあたかも自己の「手足」のように利用するとの比喩的な文脈で用いられている⁶⁾。このため、上記の判断基準における「道具」の概念も、「手足」のような限定的な意味を想定したものであるように感じられる。ところが、利用者がプロンプトを入力して画像を生成する場合、AI (人工知能) を単なる (知能のない) 「手足」として利用するのではなく、文字通り「知能」(脳の代替物) として利用することが通常である。した

* 弁護士、元東海大学実務法学研究科特任教授、東海大学総合社会科学研究所知的財産ユニット研究員、上智大学大学院応用データサイエンス学位プログラム非常勤講師

がって、上記判断基準がAIを「手足の代わり」(道具)として利用した場合を想定しているとするれば、これがAI生成物の著作物性(利用者の著作者性)を肯定するメルクマールとなり得るかとの疑問が生じる。

また、上記判断基準は、利用者の側に「創作意図」と「創作的寄与」があればAIの「道具性」が認められるものとしているが、「意図」と「寄与」という利用者側の事情のみからAI側の性質(道具性)を決定することが可能(妥当)なのか、利用者の側に「意図」と「寄与」があればAIを「道具」とみなせるとの等価性が真に成立するののかとの点についても検討が必要と思われる(後述の通り、上記判断基準には、刑法上の間接正犯(単独正犯)型の道具性のメルクマールの中に共同正犯型の寄与性の要素を混在させるのと同様の理論的課題があるように見受けられる)。

これらについて検証するには、利用者の側の事情のみでなく、利用者の脳とAIの「知能」の各機能及びそれらの相互関係について具体的に検討する必要がある。

そこで、以下においては、人間が描画表現を行う際の脳の機能について検討した後、これに続いて画像生成AIの生成過程における擬似的な「脳」の機能について検討し、そのどちらがいかなる関与の態様及び程度において創作的表現を担っていると評価できるか、それがAI生成物の著作物性(利用者の著作者性)の議論にいかなる影響を及ぼすか等の問題について考察する。

Ⅲ. 描画表現と脳

1. 総論

人間が絵を描く場合、①視覚的思考、②視覚的イメージの想起・保持及び③描画行為という3つのプロセスを一体的、重疊的、同時並行的又は相互往復的に繰り返すことが通常である。

そこで、これらがどのような脳の機能によって実現されているかについて、著作権法上の概念を踏まえながら見ていくこととする。

2. 視覚的思考 — 思想・感情、創作性(抽象的アイデア)

人間が絵を描こうとする場合、何もないところから、何を表現するために、何を対象にして描くかという問題発見型の視覚的思考と、それをどのように描くかという問題解決型の視覚的思考とが互いに重なりあい、連環となって繰り返される。この連環を生み出し、それを保持していく作用は、創造性と呼ばれる⁷⁾。

そして、これらの思考は、外界からの刺激(感覚入力)に対する反射的・受動的な反応ではなく、自発的・主体的に行動様式を選択して計画・決定する精神活動であって、前頭葉(前頭連合野)と密接な関連がある⁸⁾。特に、視覚的な美を感じるのには前頭連合野の眼窩前頭皮質(報酬に関わる領域)であるとされており⁹⁾、絵画の制作には上記領域が深く関与しているものと考えられる。

なお、描画を計画する場合、三次元的な可視世界を二次元的な像に変換して有限のスペースに配置するとの制約がある上、空間的配置自体が描画表現上の重要事項であることから、構図の決定が重要な意義を有する。そして、この思考については、後頭葉から側頭葉にかけての領域が重要な役割を担っているらしいということがわかってきている¹⁰⁾。

これを著作権法的に整理すると、何もないところから、何らかの「思想又は感情」をモチーフやテーマとして表現するために、コンセプト及び特定の視覚的な題材、対象、美観等を想定し、それを個性的な(「創作的」な)構図、画角、背景、形状、輪郭、表情、色彩、陰影、ポーズ、ディテール、画風、技法等の表現方法で描くことを企図・計画するというのが絵画の制作における視覚的思考であり、その思考は主として前頭連合野が担っており、後頭葉・側頭葉領域もこれに関連していることになる(以下、この活動を行う脳のシステム全体を「視覚的思考脳」と呼ぶ)。

しかし、この視覚的思考は未だ「アイデア」の段階にとどまっており、著作権法で保護される「表現」には至らない。

3. 視覚的イメージ — 創作的表現に結び付ける 具体的アイデア

(1) 総論

絵画を制作する場合、視覚的思考と一体的・重畳的・同時並行的・相互往復的に、作品の画像（イメージ）を頭に思い描くという作用が必然的に営まれる¹¹⁾。この視覚的イメージを作業記憶（ワーキングメモリ）に保持し続けることにより、何も書かれていない白紙の上に、頭の中に思い描いた通りの作品を表現することが可能となる。

これを著作権法的に見ると、「思想又は感情」の表現のために思案し試行錯誤を重ねる視覚的思考（抽象的アイデア）とそれに沿った描画行動（表現）とを結び付ける媒介となる具体的アイデアが作品の「視覚的イメージ」であるということになる。

視覚的イメージを形成してそれを表現行為に利用するには、対象を直接見る、脳内に保存されている視覚的記憶を呼び出す、視覚的イメージを変容させる等の方法によって視覚的イメージを想起した上、それを作業記憶（ワーキングメモリ）に保持し続ける必要がある。

(2) 対象を見る

風景や静物、人物等を写實的に描く場合、見たままの視覚的イメージを脳内に保持し、それを白紙の上に描画して表現することになる。この場合、「見る」という行為がそのまま作品のイメージに直結し、そのイメージを媒介として描画表現が行われることになる（なお、絵を描き進めるためには、作業中の絵を見ながら加筆していくことが必要であるから、絵を描く行為は「見る」という行為を必然的に伴うことになる。）¹²⁾

この「見る」という働きを担っているのは、目ではなく脳である。目は網膜に衝突する光子の刺激を受けて受動的に外界を写し出し、それを信号に変換しているにすぎず、外界を見ているのは、網膜が受け取った刺激による断片的な信号を解釈し、それらをつなぎ合わせてひとつのまとまった情報に組み立てている大脳皮質の視覚関連領域である。具体的には、網膜の異なる神経節細胞から

位置、光量、色彩（波長）についての部分的な情報が後頭葉の一次視覚野に伝えられ、そこで輪郭、動き、色などの情報がそれぞれ独立のモジュールによって分析されて視覚連合野に送られ、それらの断片的・部分的な情報がそこで統合されて像が組み立てられると考えられている¹³⁾。

さらに、その像が意味のある情報として脳に理解されるためには、後頭葉から頭頂葉・側頭葉に連なる高次連合野において、その他の感覚や記憶の情報（聴覚・触覚・味覚・嗅覚等の記憶、思い出等のエピソード記憶、知識等の意味記憶等）と統合される必要がある¹⁴⁾。すなわち、リンゴを例にとると、その手触り、味、香り、咬んだときの音などの感覚情報の記憶、過去にリンゴを食べた際の場面やリンゴに関する知識の記憶が「リンゴ」という概念で統合され、その概念が一度形成されれば、どの感覚様式の入力からでもその概念が呼び出され、それに紐づけられた様々な記憶が想起されるようになる（概念の形成・想起には前頭連合野が主として関連しているものと思われる。）。このような働きによって、脳は見ている像の意味を知ることになる¹⁵⁾。

そして、脳は、三次元空間の中の物体を自己の側からの半分しか見ていない（その意味で2.5次元的にしか世界を認識できない）にもかかわらず、裏側まで含めた三次元モデルの視覚イメージを形成することができる（このため、裏側の見えない月の形や上からは円形にしか見えない花びんの形も三次元的に理解することができる。）¹⁶⁾

しかも、実物を見ながら視覚的イメージをいま形成しているように感じていても、実際には直近の過去の視覚情報と他の過去の記憶をもとに大脳皮質の連合野が解釈して意識のワーキングメモリに保持しているものであり、厳密にはそれも過去のイメージにほかならない。

このように、脳が見ている像は、現在の現実を忠実に写し出した実像というよりも、大脳皮質の連合野が断片的・部分的な情報を統合し、様々な感覚や記憶と結び付け、自ら解釈して作り出したイメージモデル（一種の幻のようなもの）であると考えられる。

(3) 視覚的記憶を呼び出す

人間は、いま実際に見ているものを写生するだけではなく、脳内の過去の視覚的記憶を呼び出すことで視覚的イメージを想起し、そのイメージを媒介として描画表現を行うこともできる。

前述のリンゴの例で言えば、リンゴの実物が目の前になくても、「リンゴ」という言葉からその概念に紐づけられた形や色などの視覚的イメージを呼び出すことができ、この視覚的イメージに基づいてリンゴの絵を描くことができる¹⁷⁾。

なお、前述の通り、実際に見ているものを像として再構築する場合も、過去の記憶から視覚的イメージを呼び出す場合も、いずれも高次連合野の働きにより過去の記憶を含めた様々な情報を脳が解釈してイメージモデル（一種の幻のようなもの）を創出する作用という点で異ならない。

このような視覚的イメージの想起は、側頭葉から後頭葉を含む高次連合野が担っていると考えられている¹⁸⁾。

(4) 視覚的イメージを変容させる

人間（特にアーティスト）は、脳内に思い描かれた視覚的イメージを操り、それを変容させて創作活動に用いることができる。

このような作用は、脳が自らの解釈により創出した視覚的イメージモデルを前頭連合野のワーキングメモリの俎上に乗せ、これをもとに視覚的なイメージを操作して新たなイメージモデルを創出するものであるから、前述の「視覚的思考」の一種（特に「創作性」との関連の強い思考）であると思われる。

(5) 作業記憶（ワーキングメモリ）に保持する

実際の描画動作（表現行為）を行うためには、それに先行して行われた視覚的思考の結果（視覚的イメージ）を作業記憶として脳内に保存し、描画動作の過程を通じてこれを保持し続ける必要がある。これにより、描画動作という随意運動をどのように行うかとの目標の設定が可能となる¹⁹⁾。

(6) 小括

このようにして、視覚的思考（アイデア）によって企図された視覚的イメージ（三次元的イメージを二次元のキャンバス等に落とし込むイメージモデル）が脳内に形成され、作業記憶（ワーキングメモリ）の脳領域に保持されることにより、そのイメージを表現する描画動作のための手や腕の筋肉への指令が可能となる（以下、この活動を行う脳のシステム全体を「視覚的イメージ脳」と呼ぶ。）。

4. 描画行為 — 表現脳

言うまでもなく、著作権法では、脳内にとどまる視覚的思考（抽象的アイデア）や視覚的イメージ（具体的アイデア）は保護されず、それらを外部的に具現化する創作的表現が行われた場合にはじめて保護される。描画における表現行動は、若干の例外を除けば²⁰⁾、基本的に腕・手・指の随意運動による二次元平面上での絵筆・ペン等の先端の軌跡により行われる。

すなわち、前頭連合野を中心とした視覚的思考（抽象的アイデア）の結果として創出された視覚的イメージ（具体的アイデア）に従って手先の随意運動を制御し、描画表現行為を行うには、上肢、腕、手、指などの筋肉のうち、動員すべき筋肉とそうでない筋肉を選別し、動員すべき筋肉を順序よく強調的に収縮・弛緩させるタイミングを決定し、実行中の運動のモニタリングとエラーの修正を行う、との極めて精緻な計算と実行が必要となる。

この作用には、前頭葉の運動連合野（運動前野・補足運動野）、大脳基底核及び小脳が重要な役割を果たしている。すなわち、運動連合野が前頭葉の意図する動作の運動プログラムを企図し、大脳基底核がその動作に動員すべき筋肉とそうでないものを選別し、小脳が運動連合野の企図した軌道から逆算してそれを実現するのに最適な筋収縮のタイミングと強さを決定し、これらに従って前頭葉の一次運動野から手指の運動を支配する下位運動ニューロンに随意運動の指示が送られる²¹⁾。また、描画における空間認知能力は、自己の身体とそれを取り巻く三次元的世界の構造と位置と配置

を知覚し、記憶を蓄え、それに基づいて眼球と手足の運動を調節する機能を司る頭頂葉が担っている²²⁾。さらに、線を引いたり絵具を塗ったりする繊細な随意運動においては、手が目的の運動を遂行しているかどうかを確認する視覚誘導性の運動制御（フィードバック・コントロール）も必要であり、この機能は高次連合野である上頭頂小葉が司っていると考えられている²³⁾。

そして、創作的な描画を行うためには、視覚的イメージに従った動きをスムーズに行う必要があり、絵画の制作活動においては、創造性の指標のひとつに「描画の流暢性」が挙げられるという²⁴⁾。描画の流暢性にもっとも関係するのは前頭葉（特にその右半球側）と考えられており、ここが中心的な役割を担い、視覚的なイメージに関わる後頭葉・側頭葉や視覚構成機能に関わる頭頂葉の活動と相まって、新たなものを創作的に表現するという営みを司っていると考えられている²⁵⁾。

このようにして、人間の脳では、前頭葉が作り出した思想又は感情（抽象的アイデア）と創作的な視覚的イメージ（具体的アイデア）を外部的に表出させるべく、前頭葉と後頭葉・側頭葉・頭頂葉が協調的に指令を出し、運動連合野、大脳基底核、小脳、一次運動野、下位運動ニューロンなどが順次連携して最終的に手先（筆先）を動かすという創作的表現行動を発現させていることになる（以下、この活動を行う脳のシステム全体を「表現脳」と呼ぶ。）。

5. 小括

上記のように、人間が絵を描く場合、①視覚的思考（思想・感情、創作性、抽象的アイデア）、②視覚的イメージの想起と保持（具体的アイデア）、③これらに基づく描画行為（創作的表現）との各要素において、視覚的思考脳、視覚的イメージ脳、表現脳の各システムが連動しながら極めて複雑かつ精緻な活動を行なっている。

殊に、表現脳の活動とその習熟度は表現活動において極めて重要である。「描画の流暢性」が表現者の創作性（個性）の重要な要素であり、たとえば視覚的思考及びイメージの想起ができたとして

も、表現脳の習熟度が低い場合、思考及びイメージの通りの個性的表現を外部に表出することは困難だからである。

これに対し、AI利用者がプロンプトを入力して画像を生成する場合、これらの脳機能の多くが省略され、それがAIの「知能」によって代行されることになる。

そこで、次に、画像生成AIが具体的にどのような方法・機序でこれらの脳機能を代行する役割を果たしているかについて概観することとする。

IV. 画像生成AIの仕組み

1. 総論

コンピュータにより機械学習を行う人工知能（AI）は、二進法で表わされる数値を入力しこれを他の数値に変換して出力するものであり、数学的には関数である。そして、機械学習（Machine Learning）は、予測関数から損失関数を導出し、微分等により「損失関数の最小化」という数学モデルを解くことでパラメータを自動的に調整してモデルの最適化を図るものであり²⁶⁾、機械学習の一分野の深層学習（Deep Learning）に用いられるディープニューラルネットワーク（Deep Neural Network；以下「DNN」という。）は、脳の神経回路を模したニューラルネットワーク（Neural Network）を何層にも重ねることで、複雑なデータの実相に近似する関数モデルを獲得できるようにしたものである。

DNNを活用した生成モデル（generative model）は、実際に観測されたデータからなるサンプル（観測事象）がデータセットとして与えられた場合、そのサンプルの背後にある真のデータ分布（原因事象）について、結果から原因を探る逆確率を予測するベイズ推定（Bayesian inference）を応用して推論し、真のデータ分布に近似する分布モデルの確率密度関数を獲得して、その近似モデルの関数に従って未知の新たなサンプルデータ（データセットのデータと同じようなデータ）を出力できるようにした生成器（一種のシミュレーター）である²⁷⁾。

画像を出力する生成モデルには様々なものがあるが、議論の複雑化を避けるため、ここでは「プロンプト」と呼ばれる自然言語（テキスト）での指示により画像を生成する Text-to-Image の代表的なモデル (Dall-E2, Dall-E3, Midjourney, Stable Diffusion 等) の根幹をなす仕組みについて検討する。

これらの画像生成 AI は、サンプル画像を高次元（画素数等に対応する次元）のベクトルに変換して数値化し（高次元空間の中の1点を示すベクトルとして扱い）、大量のサンプル画像を学習することにより、その特徴に近似する画像を出力できるようにしたものである。上記の高次元空間（特徴空間）では、類似の特徴を有するベクトルどうしは類似度に応じた近さで連続的かつ無数に存在するものとされるため、それらの分布に近似する確率密度関数を獲得することによって、その特徴を指定された場合に近似する画像を出力することができるようになり、また特徴 A と特徴 B がプロンプトとして同時に指定された場合には、それらの特徴ベクトルの双方に近似する新たな画像を出力することができるようになる。

そして、このような Text-to-Image の画像生成モデルを実現するためには、その前提として、プロンプトとして入力されるテキストとそれに対応する画像との関係性があらかじめ学習され、その学習済みの関係性に基づくガイダンスに従って画像が生成される必要がある。

そこで、以下においては、①画像の描き方を学習する「拡散モデル」の仕組み、②画像とその特徴を表すテキストとの関係を学習した「テキストエンコーダ」の仕組み、③これらを統合した Text-to-Image モデルの理念型について、それぞれ概観することとする。

2. 拡散モデル²⁸⁾

近時、画像生成モデルの主流となっているデノイジング拡散確率モデル (DDPM: Denoizing Diffusion Probabilistic Model²⁹⁾；本稿では単に「拡散モデル」という。) は、ノイズ画像から徐々にノイズを取り去って目的の画像に至る過程を学習

する機械学習モデルである。

すなわち、拡散モデルは、物理学の非平衡熱力学 (ランジュバン動力学) を応用し、あたかも水面のインクが拡散して崩れていくように (Figure. 1 参照)、画像データ (x_0) に対して時刻 t ごとに正規分布 (ガウス分布) に従ってランダムに少しずつノイズを加えていき、完全なノイズ画像になる時刻 T までこの操作を繰り返した上 (拡散過程)、その過程を逆に巻き戻し、ノイズ画像 (x_T) が元の画像 (x_0) に戻っていく過程 (逆拡散過程) をニューラルネットワークで学習するものである。

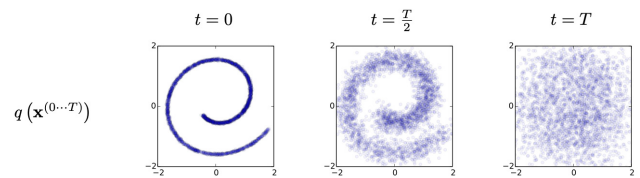


Figure. 1³⁰⁾

下図 (Figure. 2) の右から左への流れがノイズの拡散 (diffusion) を行う順過程 (forward process) であり、その反対が逆過程 (reverse process) である。

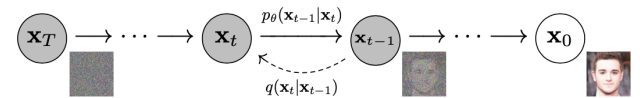


Figure. 2³¹⁾

順過程では、実際に存在するデータからサンプル画像 x_0 を取得し、そこからタイムステップごとに前の時刻 $t-1$ の画像に対して正規分布に従ったノイズ (Gaussian noise) を追加する。この過程は、条件付き確率分布 $q(x_t|x_{t-1})$ として、平均と分散によって定まる正規分布の確率密度関数により定義され、この固定された定式に従い、時刻 t における新しい画像 (時刻 $t-1$ よりわずかにノイズの多い画像) がサンプリングにより描出され、このようなプロセスが完全なノイズ画像になる時刻 T まで行われる。

画像の生成プロセスを学習するためには、上記の過程を逆に辿り、完全なノイズ画像 (x_T) から元画像 (x_0) に向けて戻る逆過程を学習することが課題となるが (Figure. 3 参照)、逆の条件付き

確率分布 $p(x_{t-1}|x_t)$ で表される逆過程自体を計算することは困難であることから、その近似分布である $p_0(x_{t-1}|x_t)$ を考える。

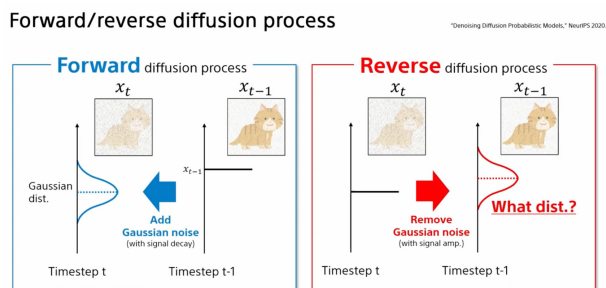


Figure. 3 ³²⁾

そして、時刻 t から $t-1$ までの微小時間の変化が十分に小さい場合には逆過程も順過程と同様の正規分布となることから、逆過程も平均と分散というパラメータにより正規分布を用いて定式化することができ、これにより追加されたノイズを確率的に予測することが可能となる (Figure. 4 参照)。

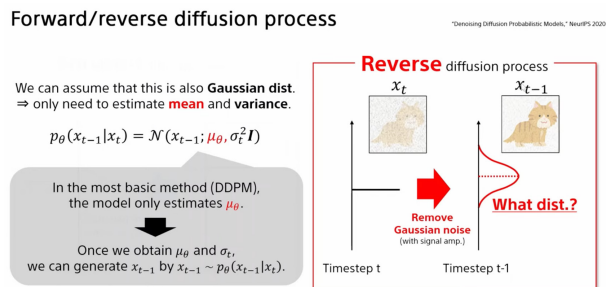


Figure. 4 ³³⁾

なお、DDPM の原論文及びその他の多くの拡散モデルでは、分散をあらかじめ固定し、条件付き確率分布の平均 μ_0 のみを学習させている。このため、一般的な拡散モデルは、タイムステップごとに U-Net その他の DNN を用いて正規分布の平均 μ_0 (パラメータ) を推定し (Figure. 4 参照)、順過程で付加されたノイズと逆過程で予測されたノイズとの間の誤差 (損失) を最小二乗法及び確率的勾配効果法を使用して最小化することをタイムステップごとに繰り返すことにより (Figure. 5 参照)、パラメータを自動的に更新してモデルを最適化し、元の画像に戻る過程の近似モデルを獲

得するものと考えられる ³⁴⁾。

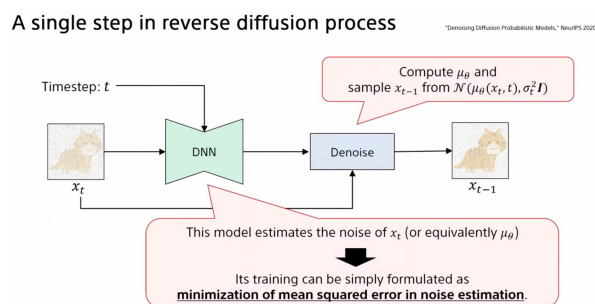


Figure. 5 ³⁵⁾

このように、拡散モデルは、白紙から始める人間の絵の描き方とは異なり、完全なノイズ画像から逆に少しずつノイズを除去 (デノイズ) して元画像に近い画像を浮か彫りにしていくという「描き方」を学習し、これにより確率的に近似する画像サンプルを生成するものと言える。

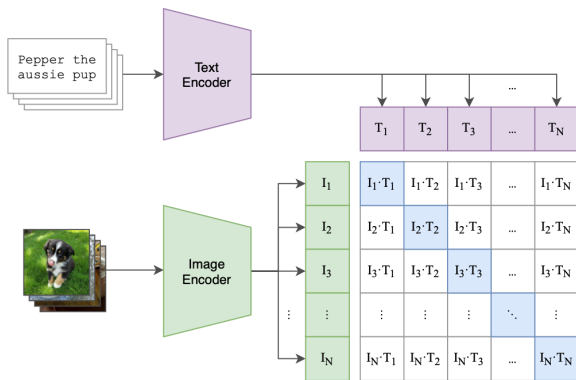
3. テキストエンコーダ

上記のように、拡散モデルは画像の描き方を学習して生成するものであるが、拡散モデル自体には、テキストで入力されたプロンプトと画像とを結び付ける機能がない。このため、Text-to-Image モデルにおいては、獲得したい画像の特徴として入力されたプロンプトを特徴ベクトルに変換し、それと類似する特徴ベクトルを導出するようにできるテキストエンコーダが必要となる。この点、Dall-E2 や Stable Diffusion などの多くの Text-to-Image モデルでは、OpenAI が開発した CLIP (Contrastive Language-Image Pre-training) と呼ばれる DNN モデルのテキストエンコーダが利用されている。

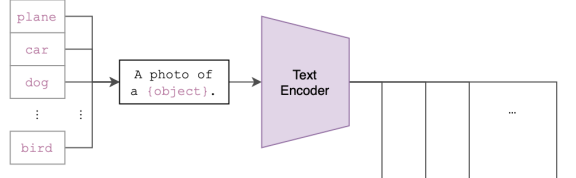
CLIP は、インターネット上から 4 億枚の画像とこれを説明した生のテキストのペアを収集してデータセットとして使用し、画像の特徴を抽出してベクトル化するイメージエンコーダ (ResNet / Vision Transformer) と、その画像に対応するテキストの特徴を抽出してベクトル化するテキストエンコーダ (Transformer ベースのモデル) とを共同で訓練し、特徴空間内の画像の特徴ベクトル (I_i) とそれとペアをなす正しいテキストの特徴ベ

クトル (T_n) のコサイン類似度を最大化する (間違った組み合わせの類似度を最小化する) ように事前対照学習 (Contrastive Pre-training) を行うことにより, 新たに与えられた画像に適したテキストを予測することができるようにしたものである (Figure. 6 参照)^{36), 37)}。これにより, CLIP は初見の画像に対してもその特徴を把握して的確なラベル付けや分類を行うことができ, 学習済み CLIP に例えば犬の写真を入力すると, 犬の画像である確率が最も高いと判定して「犬の写真」というラベルを出力する (Figure. 6(2) (3) 参照)。

(1) Contrastive pre-training



(2) Create dataset classifier from label text



(3) Use for zero-shot prediction

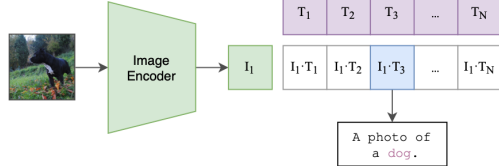


Figure. 6³⁸⁾

このように, CLIP は, 画像とその視覚的な概念を表す自然言語 (テキスト) との結び付きを学習したマルチモーダルモデルである。これを拡散モデルと組み合わせることにより, テキスト (プロンプト) で指定した特徴を有する画像を生成することが可能となる。

具体的に見ると, CLIP のテキストエンコーダ

は, 画像生成 AI に入力されたプロンプト (例えば “a cute and adorable bunny”) をエンコーダ内の Tokenizer で “a” “cute” “and” “adorable” “bunny” のようにトークン (単語等の最小単位) ごとに切り分けた上 (トークン化), エンコーダ内の Transformer でトークンごとに特徴ベクトルに変換して, その特徴ベクトルと類似するベクトルが導出されるようにガイドする。拡散モデルは, このベクトルのガイダンスに従ってデノイズを行う動作を繰り返し, ガイダンスの通りにノイズを取り去った状態のベクトルをデコーダで画像に変換することにより, プロンプトのテキストの意味に沿った画像を生成する³⁹⁾ (Figure. 7 参照)。

Extension to text-to-image generation

- Approach: add encoded text to the input of the diffusion model.

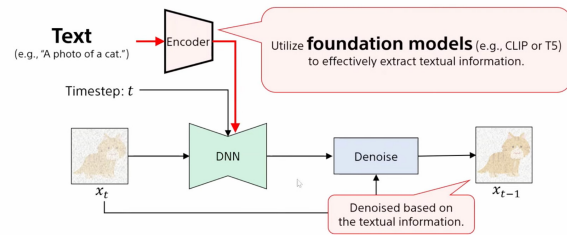


Figure. 7^{40), 41)}

4. Text-to-Image モデルの理念型

上述の通り, 生成モデル (generative model) は, 多数のデータからなるサンプルがデータセットとして与えられた場合, そのサンプルの背後にあるデータの分布を推定する学習を行い, これに近似する分布モデルを獲得して, その近似モデルの関数によって未知の新たなサンプルデータを出力するものである。

このため, 画像生成モデルの品質評価においては, 高精細な画像が生成できるかという点に加え, バリエティに富んだ新たなサンプルデータを生成できるかという点が評価される⁴²⁾。

このように, 画像生成モデルは, 画像を丸暗記して再現するというのではなく, 学習した画像群の背後にある一般的な特徴をもとに, 学習したサンプルそのものとは異なる多様なサンプルを新たに生み出す汎化能力があるからこそ高度な「知能」との評価が与えられる面があることに留意する必

要がある。

V. AI 生成画像の著作権性

1. 前提問題

前述のように、人間が絵を描く場合、①視覚的思考、②視覚的イメージ、③描画行為（表現）との各要素における脳の活動により、手と道具がコントロールされて脳内のアイデアが画像表現として具現化される。これを模式的に示したのが Figure. 8 の図である。

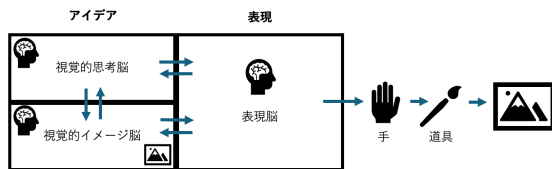


Figure. 8

これに対し、生成 AI の利用者がプロンプトを入力して画像を生成する場合の脳と AI (Generative Model) の関係を模式化すると、Figure. 9 の図のようになる。

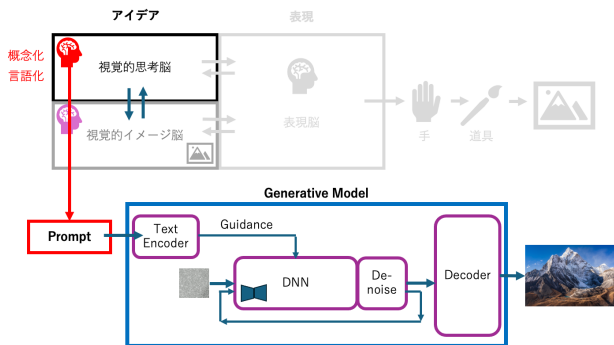


Figure. 9

上記の通り、人間が絵を描く場合、描画という表現行為を直接的に実行しているのは（手や道具を動かす）表現脳であり、視覚的思考脳及び視覚的イメージ脳は、創作意図（思想・感情、創作性）その他のアイデアを表現脳に伝えてコントロールする役割を果たしている。

ところが、利用者がプロンプトを入力して画像を生成する場合、表現行為を実行するための表現

脳は用いられない。前述の通り、「描画の流暢性」が表現者の創作性（個性）の重要な要素であり、その個性は表現脳によって外部的に表出されるものであることに鑑みれば、表現脳を用いない活動を「表現行為」と評価することには相当の慎重さが必要である。

次に、表現を導く視覚的思考脳・視覚的イメージ脳に着目しても、これらの活動と AI による画像生成との間には看過し難い断絶が見られる。前述の通り、利用者がいかに細かく視覚的思考をめぐらし、具体的で詳細な視覚的イメージを頭に思い描いたとしても、それらは概念化され言語化された抽象的なプロンプトの形を通じてしか AI には伝達されない。このため、AI は、自らが大量の文章と画像から学習した意味、特徴及び関係性（テキストと画像の結び付き）に依拠し、自らの計算に従って画像を生成するのであって、利用者の内的な視覚的思考及び視覚的イメージの多くは切断され、捨象されてしまう。例えば、利用者が「山」という言葉をプロンプトとして入力した場合、AI は自らが学習した「山」というテキストとこれに結び付けられた画像の特徴をもとに独自の計算及び出力を行うのであって、利用者の脳内画像を具現化するのではない。

また、前述の通り、生成モデルの本質は、サンプルデータとして与えられた画像群とテキスト群の背後にある一般的な特徴を学習して汎化能力を獲得し、これをもとに確率的に近似する新たな画像サンプルを生み出すものであるという点にあるため、同じプロンプトを入力したとしても、毎回同じ画像が生成されるということにはならない。

このように、生成モデルは、抽象的な言語（プロンプト）の特徴をもとに、自己の関数の確率的計算に従って多様なサンプルを新たに生み出すものであって、利用者は、プロンプトの入力等⁴³⁾の限度で AI に働きかけを行うことができるにとどまり、プロンプトの意味の解釈及びそれに基づく描画活動は AI が自らの計算で行うことになる。そして、生成 AI の各モデルは、学習してきたサンプルデータや学習方法を異にするため、それぞれに「個性」を有しており、Dall-E, Stable

Diffusion, Midjourney, Adobe Firefly 等の異なる生成モデルに全く同一のプロンプトを入力しても、同じ画像は生成されず、各モデルの描画表現上の「個性」が発揮されていると評価すべき異質な画像作品が生成される (Figure. 10 参照)。



Figure. 10 ⁴⁴⁾

このため、完成作品の詳細な視覚的イメージを先に作り上げ、デッサン、下書き、色付けなどによりイメージを詳細かつ具体的に想定した上で生成画像を得ようとするような行動は、Text-to-Image モデルの性質に照らして意味に乏しい。そのような想定が多くは、プロンプトという形で概念化し言語化した時点で捨象されてしまう上、利用者のイメージに寸分違わぬ画像が AI 独自の学習と計算から生み出される可能性は極めて低いからである。そこで、通常の利用者は、自らの欲する画像の曖昧なイメージ又は抽象的な概念のみを想定してそれを言語化したプロンプトを入力し、生成された画像が気に入らなければ次の生成の際にプロンプトを少しずつ補充・修正するなどして、好みの画像が出るまで試行を繰り返すことになる (このように、利用者から見て運任せの要素が強いという意味で、AI 画像生成は「呪文」(プロンプト)を唱えて「ガチャ⁴⁵⁾を回す」と比喩的に表現されることがある)。

このように見てみると、Text-to-Image モデルの画像生成 AI の利用者は、視覚的思考脳及び視覚的イメージ脳の活動のうちの相当部分及び表現脳の活動のほぼ全てを AI の知能に代行させ、AI

の個性的表現に結果を委ねていると評価することが可能である。

2. 基本的視座

(1) 制作委託との等質性

上記の通り、Text-to-Image モデルの生成 AI は、欲する画像の特徴を言葉で伝えることにより、人間が自ら描画表現を行う場合の脳機能を他の知能に代行させて絵を描かせていると評価できるものであるから、著作物の制作を他者に発注する「制作委託」(外部委託)と平行に考えることができる。

すなわち、画家等に絵画・イラスト等の制作を委託する場合、注文者が欲する絵の特徴をいかに詳細に言葉で伝えたとしても、成果物である美術の著作物を創作的に表現する著作者は画家であると考えられる(画家は、注文者の言語的指図を踏まえながら、自己の視覚的思考脳・イメージ脳・表現脳を駆使して美術の著作物を創作する)。

これと同様に、生成 AI の利用者が AI に描画を代行させる場合、利用者の言語的な指図(プロンプト)が詳細であったとしても、生成物である画像の著作者となるべき存在は AI であるから、結局、著作者は存在しないこととなる。

したがって、Text-to-Image モデルの画像生成 AI の利用者を生成物の著作者と解することは、基本的に困難であるように思われる。

(2) 創作意図による因果的支配性

なお、生成の過程及び結果に対して利用者の具体的な創作意図(視覚的思考脳及び視覚的イメージ脳の活動によって思想又は感情及びその創作性を具現化する意図)による因果的支配が認められる場合があり得るとすれば、利用者が美術の著作物の著作者であるとの規範的評価を行うことは理論上可能であろう。

しかし、この議論は、刑法上の間接正犯理論と同様に、因果の流れを切断する要因がなく、被利用者が単に利用者の「手足」としてのみ用いられることを前提とする⁴⁶⁾。

この点、生成モデルの本質がそれ自体の固有の

学習及び確率的な計算（推論）により新たなサンプルを創造する「知能」であることに鑑みれば、AIの裁量的表現や確率的な偶然性等の要因により利用者の因果的支配は切断されることになるであろうから、Text-to-Imageモデルの利用者が生成画像の著作者であると規範的に評価されるケースはにわかに想定し難いように思われる⁴⁷⁾。

3. 米国著作権局の見解

本論点については、米国著作権局（U.S. Copyright Office）が本稿と同旨の見解を公表している。

すなわち、同局は、2023年3月16日発効の著作権局規則「著作権登録ガイドンス」において次の通り述べ⁴⁸⁾、AI生成物の著作物性の判断はケース・バイ・ケースの判断だとしながら、プロンプトの入力により出力された生成物は人間の著作物ではないとの結論を導いている。

- ① AIが人間からプロンプトのみを受け取り、それに応じて複雑な文章的、視覚的又は音楽的作品を生み出す場合、創作の伝統的要素は、人間のユーザーではなく、AIによって決定され実行される。
- ② 現在利用可能な生成AIのテクノロジーに関する当局の理解では、ユーザーはAIがプロンプトをどのように解釈して成果物を生成するかにつき最終的な創造的コントロール（ultimate creative control）をもたない。
- ③ プロンプトは、委嘱されたアーティストへの指示のように機能する。プロンプトは利用者（プロンプター）が描きたいものを識別するが、それらの指示が出力にどのように実行されるかは機械が決定する。
- ④ AIが出力の表現要素を決定するとき、生成された成果物は人間の著作物ではない。その結果、その成果物は著作権で保護されず、その著作権登録申請は却下されなければならない。

また、同局の著作権審議会（Copyright Review Board）は、2023年9月5日付の「『Théâtre D'opéra Spatial』の登録拒否に対する再考慮の第二

次要求について」と題する回答書⁴⁹⁾において、画像生成AIのMidjourneyにテキストプロンプトを少なくとも624回入力してシーンを調整し、焦点を当てる部分を選択し、画像のトーンを指示し、画像の現像方法を指示する様々なパラメータをプロンプトに追加して「創造的な入力」を行ったと主張する出願人Jason M. Allenの下記絵画作品の著作権登録申請について、次のように述べ、却下の理由を説明している。



- ⑤ 審議会の見解では、Allen氏の上記の行為は、彼を画像の著作者とするものではない。なぜなら、彼の唯一の画像への貢献は、それを生成するテキストプロンプトを入力することだったからである。
- ⑥ Midjourneyのドキュメントによると、プロンプトはシステムが生成するものに「影響」を与え、Midjourneyによって「解釈」され、「学習データと比較」される。
- ⑦ Midjourneyはテキストプロンプトを直接の指示として扱わないため、ユーザーは満足のいく画像にたどり着くまでに何百回も反復を試行する必要がある可能性がある、というのが当局の理解である。

上記③の「委嘱されたアーティストへの指示」のアナロジーは、AIによる画像生成を著作物の「制作委託」と同視する本稿の見解と同旨であり、上記②の「最終的な創造的コントロール（ultimate creative control）」の基準は本稿の「創作的意図による因果的支配」と軌を一にするものであり、上記

①②⑤⑥⑦の AI テクノロジーとプロンプト入力に係る分析及び上記④の結論は本稿と趣旨を同じくするものであって、同局の見解はわが国の議論においても参考になる。

4. 小括

規範的評価はケース・バイ・ケースの個別判断であるから、プロンプトの入力により AI が生成した画像に著作物性が認められるケースが皆無であるとはその性質上言えないが、上記の考察により、現時点の Text-to-Image モデルの技術を前提とする限り、プロンプトを入力した利用者が生成画像を創作的に表現していると評価されるケースはにわかに想定し難いように思われる。

VI. 文化庁の判断基準の検討

1. 判断基準

文化庁は、令和5年6月、「令和5年度著作権セミナー AIと著作権」と題する資料を発表し、その中で「AI生成物は『著作物』に当たるか・著作者は誰か」との論点について以下の判断基準を示し^{50), 51)}、周知を図っている(同57~60頁)。

- ①人が思想感情を創作的に表現するための『道具』として AI を使用したものと認められれば、著作物に該当し、AI 利用者が著作者となる(命題①)
- ②人が AI を『道具』として使用したといえるか否かは、人の『創作意図』があるか、及び、人が『創作的寄与』と認められる行為を行ったか、によって判断される(命題②)
- ③創作意図(思想又は感情を、ある結果物として表現しようとする意図)は、『AI を使用して自らの個性の表れとみられる何らかの表現を有する結果物を作る』という程度の意図があれば足りる⁵²⁾(命題③)
- ④どのような行為が『創作的寄与』と認められるかについては、個々の事例に応じて判断する必要があり、生成のために AI を使用する一連の過程を総合的に評価する(命題④)

その後、文化審議会著作権分科会において、上記命題④の「創作的寄与」に関して「単なる労力にとどまらず、創作的寄与があるといえるものがどの程度積み重なっているか等を総合的に考慮して判断される」との文言が付加され⁵³⁾、その判断要素として、①指示・入力(プロンプト等)の分量・内容、②生成の試行回数、③複数の生成物からの選択が例示され⁵⁴⁾、「創作的表現といえるものを具体的に示す詳細な指示は、創作的寄与があると評価される可能性を高める」「①と組み合わせた試行、すなわち生成物を確認し指示・入力を修正しつつ試行を繰り返すといった場合には、著作物性が認められることも考えられる」「通常創作性があると考えられる行為であっても、その要素として選択行為があるものもあることから、そうした行為との関係についても考慮する必要がある」との説明が付されている⁵⁵⁾。

2. 検討

(1) 総論

文化庁の考え方は上記の通りであり、この考え方は主要な学説とも軌を一にするものであるように見受けられる⁵⁶⁾。

しかし、以下の通り、上記の考え方には、理論的にも実質論的にも検討を要する点がある。

(2) 上位規範(命題①)の検討

前述の通り、AI は道具にほかならないから、「『道具』として AI を使用したものと認められれば、著作物に該当し、AI 利用者が著作者となる」との判断基準を字義通りに捉える場合、その妥当性には疑問が生じる。

いわゆる「道具理論」は、被利用者である人間を「道具」に喩える点で意味があるが、道具そのものについて道具として使用したか否かを問題とすることは意味に乏しい。また、このような判断基準では、過度に広汎な AI 生成物が利用者の著作物と認められてしまうおそれがある⁵⁷⁾。

そこで、仮に「道具」の概念を用いるとしても、意味の限定された狭義の道具概念を定義することが必要となる。この点、著作権法上の侵害主体論

において、他者を自己の「手足」ないし「道具」として利用していると規範的に評価できる場合には、利用者が行為主体となるものとされていることからして⁵⁸⁾、前記判断基準における「道具」の概念も、主体の意思を実行するにすぎない「手足」(AI利用に関しては主として「手」)に相当するような限定的な意味を想定したものであるように感じられる。実際に、人間がペンやラケットなどの道具を操る際、脳があたかも手の一部を延長したかのような処理を行なっていることが明らかになっており、人間の脳は手と道具を一体として捉えていると考えられている⁵⁹⁾。

ところが、これまで論証した通り、AIにプロンプトを入力して画像を生成させる利用者は、AIを単なる「手」の代わりないしその一部として使用しているのではなく、視覚的思考脳・イメージ脳及び表現脳の機能を代行させる人工的な「知能」として使用している。このため、「脳の代わり」であるAIが単に「手の代わり」として使用されることはその性質上考え難く、上記の意味での道具基準は、AI生成画像の著作物性を肯定するメルクマールとしては機能し難いように思われる。

(3) 下位規範(命題②)の検討

前記の下位規範(命題②)は、利用者に何らかの「創作意図」と「創作的寄与」があればAIの「道具性」が認められるとするが、これにも疑問がある。

まず、「意図」と「寄与」という利用者側の事情のみからAIの側の性質を決定すること自体に疑問がある。AIの道具性を認定評価するためには、利用者側の事情のみならず、AI側の事情(当該AIの性質・機能等)及びAIと利用者との関係性を考慮する必要がある(本稿Ⅲ～Ⅴ参照)。

また、利用者側に「創作意図」と「創作的寄与」があることと、AIが単なる「道具」であることは、等価とは言い難い。いかに人間側に「意図」と「寄与」があろうとも、AIの知能活動により当該AI固有の画像が生成される以上、AIを「手足」のような「道具」と評価することは難しい。利用者側に「創作意図」と「創作的寄与」があることは、

利用者を著作者と評価するために最低限必要な要素にすぎず、それで足りることにはならない。

そもそもこの議論の問題性は、刑法上の間接正犯(単独正犯)に類する「道具」の概念を援用しながら、そのメルクマールにおいて共同正犯に類する「寄与」の概念を援用している点にあるように思われる。前者は、因果を支配する利用者を単独正犯と評価するものであり、後者は、意を通じて互いに他人の行為を利用し合う複数の関与者どうしを共同正犯とするものであって、両者の趣旨及び要件は異なる。利用される側が「道具」であることを前提に利用者側の主体性を議論する以上、間接正犯理論と同様、利用者が自己の意思に基づき事態の成りゆきを因果的に「支配」していることが必要であり、利用者が単に「寄与」しているというだけでは足りない。前記の下位規範(命題②)は、単独著作について間接正犯(単独正犯)に類する道具理論を援用しながら、共同著作(共同正犯)に類する「寄与」のメルクマールを設定している点で、理論的に飛躍があるように思われる。

さらに、創作の「意図」と「寄与」がありさえすればAI利用者の著作者性が肯定されるとの価値判断は、具体的妥当性の観点からしても広きに失するものと考えられる。人間が視覚的思考脳・イメージ脳及び表現脳を駆使して自ら創作的表現を行った場合にはじめて単独で著作者と認められることとのバランスに鑑みれば、著作者性を肯定するには、単に「寄与」したというだけでは足りず、自己の具体的な創作意図による因果的な「支配」が必要と解すべきである。

(4) 命題③の検討

前記命題③は、「創作意図」について、『AIを使用して自らの個性の表れとみられる何らかの表現を有する結果物を作る』という程度の意図があれば足りる」としており、これを支持する学説もある⁶⁰⁾。

しかし、著作物は「思想又は感情を創作的に表現したもの」であり、(美術の著作物について言えば)前述の通り視覚的思考脳及び視覚的イメージ脳の活動により、自らが表現したい思想又は感

情を創作的に表出するための「アイデア」段階の内部的な意図が具体的に形成され、それが表現脳の活動によって具現化されることにより外部的に成立することとなるのであるから、その創作意図は、最終的な著作物を具現化するに足る具体性を具備している必要があり、「何らかの表現を有する結果物を作る」との意図では足りないように思われる（なお、創作意図による因果的支配の有無を規範的評価の基準とする本稿の立場からは、その創作意図は最終的な表現を支配するに足る具体性を具備したものでなければならないとの結論が自然に導かれる。）。

(5) 命題④の具体化規範の検討

文化庁は、「創作的寄与」の判断における総合的考慮の要素として、①指示・入力（プロンプト等）の分量・内容、②生成の試行回数、③複数の生成物からの選択を例示し、「創作的表現といえるものを具体的に示す詳細な指示は、創作的寄与があると評価される可能性を高める」と述べるが、この説明を額面通り読む場合、具体的かつ詳細なプロンプトを入力し、生成物を確認しながら指示・入力を修正しつつ試行を繰り返して画像を選択した利用者の多くは、生成画像の著作者として著作権を享受するとの期待を抱くことになると推察される（実際に、学説の中には、このような場合に著作者性を認める見解がある⁶¹⁾）。

しかし、前述の通り、現時点での技術を前提とする限り、詳細にプロンプトを入力した場合であっても、利用者が著作者となるケースはにわかには想定し難いと思われることから、利用者に過度の期待を抱かせることのない慎重な表現が望まれる。

(6) 共同著作物のアナロジーの妥当性

文化庁の判断基準は、立法的措置の存在しない現時点において、共同著作物における議論を参考にしながら、現行法の解釈によるAI生成画像の保護の余地を残しておくために工夫されたものであると考えられ、「創作的寄与」との文言も、共同著作物に係る「2人以上の者が共同して創作し

た著作物であって、その各人の寄与を分離して個別的に利用することができないもの」（法2条1項12号）との法文を参考にしたものと推察される⁶²⁾。実際に、文化審議会著作権分科会の資料においても、「共同著作物に関する裁判例等⁶³⁾に照らせば、生成AIに対する指示が表現に至らないアイデアにとどまるような場合には、当該AI生成物に著作物性は認められない」と述べられており⁶⁴⁾、生成AIの著作物性の議論において共同著作物への言及が見られる。

しかし、人間と機械（AI）との共同著作が観念できないことは言うまでもなく、これを認めるに等しい解釈を採る場合、現行法の体系との整合性に疑問が生じる。

また、ここで問題となっているのは、生成AI利用者の単独での著作者性であり、権利行使等に制限のある共同著作物の著作者と同様の要件によって単独著作者としての地位を認める場合、実定法を上回る過剰な保護となりかねない⁶⁵⁾。

3. 小括

以上の観点から、Text-to-Imageモデルの画像生成AIに関する考察から検討する限り、文化庁の整理する判断基準には再検討の余地があるように思われる。

VII. その他の論点

1. 偶然性に関する議論

学説の中には、成果物の出来具合が偶然性に左右される場合にも著作物の成立を肯定できるとの前提に立ち、AI生成物の著作物性（利用者の著作者性）を広く認める見解がある^{66), 67)}。

しかし、AIは学習した膨大なサンプルに従って概念をベクトル化し、獲得した関数により計算（推論）を行って結果を出力しているものであり、生成画像に生じる差異は単なる偶然の産物ではなく、各AIの知能（学習と推論）に基づく「個性」によって生じる側面が大きいから、偶然性の介在する著作物の存在を示すことによってAI生成物の著作物性を肯定する議論には飛躍があるように

思われる。

また、AI生成物の著作物性は規範的評価の問題であり、偶然性の介在は規範的評価の一要素にすぎないから、この要素に係る側面のみから著作物性（著作者性）を肯定する議論を行うことは難しい。

この点、本稿の立場からは、利用者の具体的な創作意図による因果的支配があると規範的に評価できるかどうかとの判断に尽きることになるから、偶然性の介在もその一要素となり、それが一定限度を超えれば因果的支配が切断されると評価すべきことになるが⁶⁸⁾、因果的支配の切断を生じさせるものは偶然性の要素だけではなく、他の知能（人間、AI）による固有の解釈や裁量等の介在も因果の切断を生じさせる要素となる。したがって、本稿の立場からは、これらの諸要素を総合的に斟酌し、それでもなお利用者の創作意図による因果的支配があったと評価できれば著作物性（利用者の著作者性）を肯定すべきことになるが、偶然性の要素が介在する著作物もあるというだけの理由でAI生成物の著作物性を肯定することはできない^{69), 70)}。

2. 写真著作物との比較

AI生成物の著作物性を検討するにあたっては、写真の著作物と比較することが有用である。機械の自動機能が関与して結果物がアウトプットされるという意味で生成AIと共通の要素がある反面、それらの利用のされ方には本質的な相違があるからである。

すなわち、写真は、被写体の選択・組合せ・配置、構図・カメラアングルの設定、シャッターチャンス捕捉、被写体と光線との関係（順光、逆光、斜光等）、陰影の付け方、色彩の配合、部分の強調・省略、背景及びそのボケ具合等の諸要素を総合してなる一つの表現であり、上記の諸要素のうち一部が機械によって自動的に調整されるとしても、それらの重要部分について撮影者が「視覚的思考」を働かせ、カメラのファインダーを覗くなどして形成した「視覚的イメージ」を有形的に固定することが通常であるから、結果物に対する撮影者の

創作意図による因果的支配が認められる場合が多い。この点、知的財産高等裁判所の判例は、広告販売用のホームページに掲載されていた商品写真について、「被写体の組合せ・配置、構図・カメラアングル、光線・陰影、背景等にそれなりの独自性が表れている」と述べて著作物性を認めながらも、「その創作性の程度は極めて低いものであって、著作物性を肯定し得る限界事例に近いものといわざるを得ない。」と判示している（知財高判平成18年3月29日〔平成17年（ネ）第10094号スメルゲット事件〕）。

これに対し、画像生成AIは、前述の通り、利用者の「視覚的イメージ」から切り離された当該AI固有の解釈並びに学習及び推論（関数計算）に基づいて結果物を出力するものである。

したがって、撮影者の視覚的思考により形成された視覚的イメージを具現化した表現物である商品写真が「著作物性を肯定し得る限界事例」であるとする、利用者の視覚的イメージから切断される形でAIにより独自に生成される画像について著作物性を肯定することは極めて困難との前述の結論が自然に導かれる。

3. 編集著作物性・二次的著作物性

前述の通り、画像生成AIは、学習用の画像データ自体を保存しているわけではなく、画像データの特徴を学習して関数のパラメータを獲得し、その関数に従って新たに近似的なサンプルデータを出力するものであって、元の画像を選択して組み合わせているものではない（生成画像は元の画像データの選択又は配列の結果ではなく、編集著作物に関する議論は画像生成AIにはなじまない）。

また、前述の通り、Text-to-Imageモデルの生成画像が著作物と評価されることはにわかに考え難いことから、それがプロンプト（言語著作物）の二次的著作物となることもまた考え難い⁷¹⁾。

VIII. 結論

以上述べてきた通り、Text-to-Imageモデルが生成した画像について、利用者の具体的な創作意

図による因果的支配が認められる場合があるとすれば、その利用者を生成画像（美術著作物）の著作者と規範的に評価することは理論上可能であるが、描画する脳の機能と画像生成 AI の「知能」としての機能及びそれらの関係性からして、膨大なサンプル学習によって獲得した膨大なパラメータを有する関数により固有の確率的計算（推論）を行って個性的な画像を生成する現時点の Text-to-Image モデルの技術を前提とする限り、利用者を画像の著作者と評価することは基本的に困難であり、現実的にもそのような規範的評価がなされる事例はにわかに想定し難い、というのが本稿の結論である。

IX. 結語

これまで述べてきた通り、画像生成 AI の生成物を現行の著作権制度で保護することについては懐疑的に考えざるを得ず、政策的な観点から解釈論によって著作権の枠組みの中での保護を広げる場合には、著作権法体系の整合性に歪みが生じることが危惧される⁷²⁾。

また、画像生成 AI の利用行為の本質は、他人の著作物を大量に含む既存の画像の特徴を学習して技術を磨いた画家に描画を委託するに等しいものであり、基本的に自己の視覚的思考脳及び表現脳の活動を大幅に省略し、元の画像の存在と描画者のクリエイティビティに依拠して、いわば呪文を唱えて好みの画像が機械から出てくるのを待つというに近い行動類型であって、実質論的に見ても、著作権法の価値体系においてさほど保護価値が高いものであるようには見受けられない。

そして、AI を「道具」として使用すれば著作権が得られ、創作面での「何らかの意図」と「寄与」があれば「道具」として使用したことになり、その周知を行う場合、一般の利用者に誤解と過度の期待を与えることにもなりかねない。

他方で、プロンプトの「アイデア」自体が個性的であり、それによって（それまでこの世になかったような）画像が生成された場合、その「アイデア」と AI 技術の融合によって得られた画像に何

らの保護も与えなくてよいとの結論には、やや違和感を覚えることも事実である。

また、プロンプトの工夫におけるノウハウは日々進化しており、その努力、労力、投資、そこから得られる営業上の利益等について、競争法的な保護を与えるとの政策判断はあり得る。

さらに、AI 生成画像である限りどのような剽窃（パクリ、ただ乗り）も許されるとの価値観が広まることは、健全な法的バランス感覚の維持・醸成という観点からも好ましいことではない。

このような立法事実が認められるとすれば、解釈論の模索により現実的な保護を遅らせるよりも、保護価値の比較的高い一定の類型につき、保護価値に応じた過不足のない立法的解決を迅速に検討することが相当という考え方もあり得る。

その場合、①著作権法の中で権利として保護するか、不正競争防止法のように行為を規制することによって反射的に保護するか、②保護要件をどのように設定するか、との2点が重要な問題となるように思われる。

現時点における筆者の雑感としては、上記①の立法類型について、「アイデア」を著作権法の中で保護することには違和感があること、競争法的な観点から「パクリ」「ただ乗り」に一定の制限をかけることが主たる立法趣旨と考えられること、不法行為の特別類型を立法化することにより、解釈論による一般不法行為の成否の議論を乗り越えることが可能であること等の理由から、行為規制型の立法によることが相当ではないかとの意見にやや傾いている。

また、上記②の保護要件（規制要件）については、①ある程度の周知性もしくはパブリシティ権類似の顧客吸引力又はアイデアの独自性その他保護価値の程度に係る要件を定めるか、② AI 生成物であること及びその生成過程の各立証（特に入力したプロンプトの証明）をどの程度求めるか、③規制行為の類型として、デッドコピー又はこれと同視し得る行為のみを禁じるか、④マーケットでの競合（混同）等を要件とするか、⑤保護期間をどの程度に設定するか等が議論の対象になるように思われる。これらの点について検討することは本

稿の範疇を越えるが、AIを利用した画像生成行為そのものが、他者の視覚的思考、表現上のアイデア、工夫、技術、労力、投資等を勞せずして簡易に利用する面があるものであることから、AI生成画像に対する保護要件については相応の絞りをかける必要があるように感じている。

解釈論によって著作権という強力な権利をAI生成画像に認めることに懐疑的な本稿の立場としては、政策的な保護価値に比例するきめ細かな立法措置の具体的検討が行われることを期待したい⁷³⁾。

(注)

- 1) 現行法での著作物としての保護に否定的な見解として、内田剛「コンピュータ生成作品の著作権による保護とその保護のための課題—オーストラリアにおける3つの判決からの示唆—」渋谷達紀教授追悼論文集『知的財産法研究の輪』554頁(発明推進協会, 2016)、現行法の解釈によるAI生成物の著作権保護に懐疑的な見地から問題提起を行うものとして、上野達弘「人工知能と機械学習をめぐる著作権法上の課題——日本とヨーロッパにおける近時の動向」(法律時報91巻8号33~40頁)、偶然性に左右される度合いが大きい成果物の著作権保護を否定する見解として、田村善之「著作権法概説[第2版]」398~399頁(有斐閣, 2001年)、具体的かつ詳細な指示などの創作的寄与がある場合には偶然性に左右されても現行法での著作権保護を肯定できるとする見解として、愛知靖之「AI生成物・機械学習と著作権法」(パテント, 2020, Vol. 73 No. 8 (別冊No.23) 131~134頁)、AIを道具として創作したものが多くことからAI生成物の多くが現行法で保護されるとする見解として、横山久芳「AIに関する著作権法・特許法上の問題」(法律時報91巻8号50~56頁)、人が「道具」としてAIを使用した場合には著作物性(利用者の著作物性)が認められ、AIを「道具」として使用したといえるか否かは「創作意図」と「創作的寄与」の有無によって判断されるとの文化庁の枠組みに賛成した上、「最終的な結果物の表現をどこまで意図的にコントロールしていたかは問われない」との見地からAI生成物の著作物性を広く肯定する見解として、前田健「AI生成物の著作物性」(上野達弘・奥邨弘司編著『AIと著作権』(勁草書房, 2024) 150~175頁(第7章))等がある。
- 2) AI利用者(「AI絵師」等)とイラストレーターらの意見の対立が指摘されている。
- 3) 文化庁「令和5年度著作権セミナー AIと著作権」(令和5年6月) 57~60頁
- 4) 文化庁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) 39~40頁「(3)生成物の著作物性について」
- 5) 広辞苑第七版「道具」「物を作り、また事を行うのに用いる器具の総称」
- 6) 中山信弘「著作権法[第4版]」(有斐閣) 763頁、作花文雄「詳解著作権法[第6版]」(ぎょうせい) 643頁、上野達弘「著作権法における『間接侵害』」(ジュリストNo.1326 [2007.1.1-15]) 77頁
- 7) 岩田誠「増補新装版 見る脳・描く脳 絵画のニューロサイエンス」(2018年, 東京大学出版会) 178~180頁
- 8) 前掲注7) 岩田 187~190頁参照
- 9) 石山徹・田中彰夫・池田るり子「描画学習における機能的な展開可能性に関する研究」美術教育学(美術科教育学会誌) 第35号(2014年3月) 84頁
- 10) 前掲注7) 岩田 114~118頁
- 11) 前掲注7) 岩田 64~71頁参照
- 12) 前掲注7) 岩田 12~13頁
- 13) 前掲注7) 岩田 27~30頁, 76~78頁。なお、同書66頁には「視覚的イメージといわれているものは、この作業記憶のなかの視覚的・空間的な部分であると思われる。」との記述があることから、像を見ていると自覚しているのは、前頭連合野の中のワーキングメモリ(作業記憶)を司る領域であると思われる。
- 14) 前掲注7) 岩田 27~30頁
- 15) 前掲注7) 岩田 42~46頁
- 16) 前掲注7) 岩田 137~153頁
- 17) 前掲注7) 岩田 64~66頁
- 18) 前掲注7) 岩田 66~71頁
- 19) 前掲注7) 岩田 90~94頁
- 20) 口や足などを用いて絵が描かれることもあるが、その場合の口や足は手の代替と考えることができる。
- 21) 前掲注7) 岩田 80~91頁
- 22) 前掲注9) 石山ら 81頁
- 23) 前掲注7) 岩田 80~91頁
- 24) 前掲注7) 岩田 187~190頁
- 25) 同上
- 26) 赤石雅典「最短コースでわかるディープラーニングの数学」(日経BP社, 2019) 4頁
- 27) C. Luo; Understanding Diffusion Models: A Unified Perspective (arXiv: 2208.11970, 2022)、岡野原大輔「ディープラーニングを支える技術2 ニューラルネットワーク最大の謎」(技術評論社,

- 2022) 83~88 頁
- 28) J. Sohl-Dickstein et al.; Deep Unsupervised Learning using Nonequilibrium Thermodynamics (arXiv: 1503.03585, 2015), J. Ho et al.; Denoising Diffusion Probabilistic Models (arXiv: 2006.11239, 2020), 前掲注 27) Luo, 岡野原, 岡野原大輔「拡散モデル データ生成技術の数理」(岩波書店, 2023) 39~42 頁, N. Rogge et al.; The Annotated Diffusion Model (<https://huggingface.co/blog/annotated-diffusion>, 2022), Masato Ishii; Diffusion models with/as Foundation Models (SONY nnabla Deep Learning Channel, https://www.youtube.com/watch?v=HfTD5_gZX4, 2022), 早川顕生「【AI 論文解説】物理学の知識を背景とした画像生成手法 Part1 Diffusion Probabilistic Models」(SONY nnabla Deep Learning Channel, https://www.youtube.com/watch?v=DDGgKt_CyRQ&list=PLbtqZvaoOVPCqfmnrBfo9Xv5mtDr0LjQZ&index=16, 2021)
- 29) 前掲注 28) Ho
- 30) 前掲注 28) Sohl-Dickstein, Figure.1 から抜粋して転載
- 31) 前掲注 28) Ho, Figure.2 から転載引用
- 32) 前掲注 28) Ishii 資料から転載引用
- 33) 同上
- 34) なお, 物理学の非平衡熱力学 (ランジュバン動力学) のアナロジーによるガウス分布のランダムノイズを用いずとも, Blur (ぼかし), Mask (マスキング), Pixelate (モザイク化) 等のランダム性のない決定論的 (deterministic) な画像変換の逆過程を学習することでも同様の効果が得られることがわかってきている。A. Bansal et al.; Cold Diffusion: Inverting Arbitrary Image Transforms Without Noise (arXiv: 2208.09392, 2022)
- 35) 前掲注 28) Ishii 資料から転載引用
- 36) A. Radford et al.; Learning Transferable Visual Models From Natural Language Supervision (arXiv: 2103.00020, 2021) p.2
- 37) OpenAI; CLIP: Connecting text and images (<https://openai.com/research/clip>)
- 38) 前掲注 36) Radford から転載引用
- 39) S. Lee et al.; Diffusion Explainer: Visual Explanation for Text-to-image Stable Diffusion (arXiv: 2305.03509v2, 2023, <https://doi.org/10.48550/arXiv.2305.03509>), S. Lee et al.; Diffusion Explainer, <https://arxiv.org/pdf/2305.03509.pdf>
- 40) 前掲注 28) Ishii 資料から転載引用
- 41) なお, Stable Diffusion では, パフォーマンスと計算効率のバランスを図る観点から, プロンプトに入力できるトークン数は 77 個に限定されている (トークンのうち冒頭と末尾は開始と終了を示す <start> と <end> であるため, 内容的なトークンは 75 個に限定される。前掲注 39) Lee 参照)。
- 42) 最もよく用いられる評価方法である Inception Score と Frechet Inception Distance (FID) は, いずれも高品質なデータが多様に生成できるかとの点を評価するものである。SONY nnabla Deep Learning Channel (YouTube) 「【Deep Learning 研修 (発展)】データ生成・変換のための機械学習 第 1 回後編『イントロダクション』」参照
- 43) ハイパーパラメータの設定も利用者が行い得る働きかけのひとつであるが, 画像作品の核心的な諸要素を決定するのはプロンプトである。
- 44) Figure. 10 は, 筆者が Open AI の Dall-E3 と Adobe の Firefly のそれぞれに同一のプロンプト (下記 ①②) を入力して得た 2 種類の画像群を示したものである (恣意性を排除するため, 各プロンプトを入力して最初に出力された画像をそのまま用いている。なお, Firefly では 4 画像が同時に出力されるが, 要部の破損や指定要素の欠損があるものは排除している)。上記の通り, 同一プロンプトの 3 画像はそれぞれ異なる作品であり, 特に Dall-E3 作品と Firefly 作品の間には各生成モデルの「個性」による相違が明確に見られる。このため, 同一のプロンプト (例えば②) を筆者が入力したことをもって, 筆者が上記 3 作品全て (例えば Dall-E3 ②と Firefly ② a, ② b) の美術著作物の著作者となると評価することは困難と思われる (このうち 1 枚についてのみ筆者が著作者となると考えることは, 同一プロンプトの入力行為を「表現行為」と解する場合, より困難であろう)。
- ① A small Japanese Shiba dog sitting on the night seashore near the Arctic Ocean, looking up at the sky and shedding tears, the big moon and the northern lights (aurora) shining in the sky, white steep mountains to the left, an old Swedish style wooden house to the right, realistic photo style image
- ② A pretty 25-year-old Japanese lady in a flower-painted pink and white kimono, standing on the seashore of Hokkaido, along with a Japanese Shiba dog by her right side, looking up at the evening sky of the beautiful sunset time, white mountains and cherry blossoms in full bloom to her left, an old Japanese style wooden house to her right, realistic photo style image
- 45) 何が出てくるかわからないままハンドルを回して

- アイテム（カプセルトイ）を獲得するシステム又はそこから転じてソーシャルゲームなどで課金してくじ引きのようにアイテムを獲得するシステム。近時では「親ガチャ」（子が親を選べず親が運で決まること）等の造語にも用いられる周知表現になっている。
- 46) 間接正犯について、藤木英雄「新版刑法」（弘文堂）140～141頁には次の記述がある。「自己の意思に基づき、自己の行動から生ずる事態の成りゆきを思いのままに操作し、それによって所期の犯罪事実実現の目的を遂げた者は、みずから手を下して実行したものと同視してよい。この見地から、他人を道具のごとく利用し自分の犯罪意思を実現する者もまた正犯であるといつてよい。」「自己の犯罪意思を実現するためにみずから事態の成りゆきを操作した者（行為支配）を正犯として責任を負うべき者とするのが、事理に即した正犯観である」]
- 47) 生成 AI 技術の進歩や利用者側の技術の発展等によって、利用者の脳内イメージが生成されたと評価されるケースが現れる可能性がないとも限らないことから、理論上の判断基準を明確にしておく意味はあると思われるが、現時点においては、限局的な特殊事例の発生に備えるためのものと位置付けることが相当ではないかと思われる。
- 48) U.S. Copyright Office, Library of Congress; Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence (A Rule by the Copyright Office, Library of Congress on 03/16/2023), “III. The Office’s Application of the Human Authorship Requirement”
<https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence>
- 49) The Review Board of the United States Copyright Office; Re: Second Request for Reconsideration for Refusal to Register Théâtre D’opéra Spatial (SR # 1-11743923581; Correspondence ID: 1-5T5320R), <https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf>
- 50) 当該判断基準は、文化庁「著作権審議会第9小委員会（コンピュータ創作物関係）報告書」（平成5年）第3章I. 1及び知的財産戦略本部「新たな情報財検討委員会報告書」（平成29年）36頁の整理を踏襲したものである。
- 51) 文化審議会著作権分科会法制度小委員会令和5年度第1回会議（同年7月6日）資料3「AIと著作権に関する論点整理について」5頁も同様の整理を行なっている。
- 52) 文化庁「令和5年度著作権セミナー AIと著作権」（令和5年6月）58頁「創作意図は、生成のためにAIを使用する事実行為から通常推認しうるものであり、また、具体的な結果物の態様についてあらかじめ確定的な意図を有することまでは要求されず、当初の段階では、『AIを使用して自らの個性の表れとみられる何らかの表現を有する結果物を作る』という程度の意図があれば足りると考えられます。』
- 53) 同法制度小委員会令和5年度第5回会議（同年12月20日）資料「AIと著作権に関する考え方について（素案）」
- 54) 第4回会議（同年11月20日）資料1-1「AIと著作権に関する考え方について（骨子案）」9頁では、「生成の際に、生成 AI に対してどの程度具体的な指示を与えれば、生成物に著作物性が認められるのか。以下のような要素は著作物性の有無に関して、生成物のどの範囲に、どの程度影響するか。他に影響が考えられる要素はあるか。①指示・入力（プロンプト等）の分量・内容 ②生成の試行回数 ③複数の生成物からの選択 ④生成後の加筆・修正」との整理がなされていたが、同第6回会議（令和6年1月15日）の資料1-2及び同第7回会議（同年2月29日）の資料2-2では、上記の要素④（生成後の加筆・修正）が削除されている。
- 55) 同年3月15日に同小委員会が公表した「AIと著作権に関する考え方について」39～40頁「(3)生成物の著作物性について」の「イ生成 AI に対する指示の具体性と AI 生成物の著作物性との関係について」
- 56) 小泉直樹ら著「条解著作権法」（弘文堂、2023）33～34頁「(3)AI生成物と『著作者』（愛知靖之執筆部分）には次の記述がある。「人間が動物やコンピューター（AI）を『道具』として用いてでき上がった表現が著作物であり、それを作出した者が著作者たり得ることは言うまでもない。」「創作的表現それ自体を作出する行為を自ら行ったと評価できるか否かを、『創作意図』と『創作的寄与』という観点から、ケース・バイ・ケースで判断せざるを得ないであろう。」「『学習済みモデル』の利用者に創作意図があり、極めて具体的かつ詳細な指示等を行った場合には、創作的寄与を肯定し、AI利用者自身がAIを『道具』として用いて創作的表現それ自体を作出する行為を自ら行ったと評価することは可能だと思われる。」なお、前掲注1) 愛知133頁及び同上野・奥邨150～175頁（前田

- 健執筆部分)も同旨。
- 57) 実際に、前掲注1) 横山には「現時点では、AI生成物といっても、人がAIを道具として創作したものが多くと考えられるため、AI生成物の多くは現行法上保護されることになるであろう」との記述がある。
- 58) 中山信弘「著作権法〔第4版〕」(有斐閣)763頁、作花文雄「詳解著作権法〔第6版〕」(ぎょうせい)643頁。なお上野達弘「著作権法における『間接侵害』」(ジュリストNo.1326〔2007.1.1-15〕)77頁ではこの考え方は端的に「手足論」と表現されている。
- 59) L. E. Miller et al.; Sensing with tools extends somatosensory processing beyond the body (Nature, volume 561, p239-242, 2018), L. E. Miller et al.; Somatosensory Cortex Efficiently Processes Touch Located Beyond the Body (Current Biology 29, 4276-4283, 2019, <https://doi.org/10.1016/j.cub.2019.10.043>)
- 60) 前掲注1) 上野・奥邨156~157頁(前田健執筆部分)「『創作意図』は、ある具体的な結果物を作成しようという意図が認められれば足り、その結果物の中身に対して具体的な意図を有している必要はないという立場を支持したい。なぜなら、創作意図を求める趣旨は、人間の創作物とはおよそ評価できないもの、すなわち、人間が全く意図せずに偶然にできあがってしまったもの…を著作権の保護対象から除くこと、及び、概括的に表現を作成する場や道具を用意したにすぎない者に対する保護を否定することに止まると解されるからである。この趣旨からすると、著作者たるべき者が、最終的な結果物の表現をどこまで意図的にコントロールしていたかは問われないこととなり、結果物がある程度偶然に左右される場合であっても、著作者と認めることができる場合はあるというべきである。」
- 61) 前掲注56) 条解34頁「(3)AI生成物と『著作者』(愛知靖之執筆部分)「『学習済みモデル』の利用者に創作意図があり、極めて具体的かつ詳細な指示等を行った場合には、創作的寄与を肯定し、AI利用者自身がAIを『道具』として用いて創作的表現それ自体を作出する行為を自ら行ったと評価することは可能だと思われる。」
- 62) 著作権法上、「創作」的な「寄与」との文言が用いられているのは、共同著作物に係る同法2条1項12号と映画の著作物に係る同法16条本文のみであり(旧法に係る附則10条を除く。)、これらはいずれも複数人の関与を前提とする規律である。
- 63) 最判平成5年3月30日判時1461号3頁〔智恵子抄事件〕、大阪地判平成4年8月27日判時1444号134頁〔静かな焔事件〕等
- 64) 同法制度小委員会令和5年度第4回会議(同年11月20日)資料1-1「AIと著作権に関する考え方について(骨子案)」
- 65) 映画著作物の著作者に係る「全体的形成に創作的に寄与した者」との要件についても、多数人の関与する映画の特殊性に鑑みて設けられた特別規定であり、これをAI生成物の著作者性の要件として一般化することは相当と思われない。
- 66) 前掲注1) 愛知133頁「目隠しをした状態で、何色かもわからないカラーボールをいくつもキャンバスに投げつけて作品を制作するという手段など、むしろ偶然性を活用した表現が行われる場合、あえて著作物性を否定することが妥当なのであろうか。」「具体的かつ詳細な指示など『人間の創作活動』に相当する寄与がある場合には、たとえその指示により成果物の出来具合が意図通り実現できるのかがブラックボックスであり、成果物の生成が偶然性に左右されるという側面があったとしても、そのような指示等を行った者を著作者として著作物の成立を認めて差し支えないと思われる。」
- 67) 前掲注1) 上野・奥邨157頁(前田健執筆部分)「創作意図を求める趣旨は、人間の創作物とはおよそ評価できないもの、すなわち、人間が全く意図せずに偶然にできあがってしまったもの…を著作権の保護対象から除くこと、及び、概括的に表現を作成する場や道具を用意したにすぎない者に対する保護を否定することに止まると解される」「この趣旨からすると、著作者たるべき者が、最終的な結果物の表現をどこまで意図的にコントロールしていたかは問われないこととなり、結果物がある程度偶然に左右される場合であっても、著作者と認めることができる場合はあるというべきである。」「そもそも伝統的な著作物においても、著作者と認められるために、結果物の創作的表現を予め意図的にコントロールすることは求められてこなかったものと思われる。美術工芸品として美術の著作物に該当するとされている陶芸についても、窯の炎などの偶然性に左右されて作品ができあがる点は、著作物たり得ない根拠とはされてこなかった。」
- 68) 偶然性は程度の問題であり、全体的な規範的評価の一要素にすぎない。例えば、毛筆による「書」について見ても、筆の毛の一本一本が辿る軌跡(毛

先のばらけ方やかすれ方)は一定の偶然性を必然的に伴うが、創作意図の因果的支配との関係において、この偶然性は無視できると規範的に判断されるであろう(全体として見れば、書家が「書」を因果的にコントロールしていると規範的に評価される。)

- 69) 前掲注 66) の見解は、目隠しをした状態でカラーボールを投げてできる結果物に著作物性を認めることを前提として AI 生成物の著作物性を肯定する結論を導くものであるが、この前提自体にも難しい問題が含まれている。色のわからないボールを投げてできる跡の著作物性の判断は個別具体的な評価の問題であり、一般化することが難しい(上記の設例も、どこまで偶然性の要素の大きなものであるかが必ずしも明らかではない。)。例えば、ボールの大きさ、材質、硬さ、薄さ、割れ易さ、絵の具の色とその個数・配分、ボールの投げ方(速度、方向、角度、密度)等の諸要素を自ら決定して投げる場合、目隠しをしていても、結果物の全体像(ボールが割れてできる絵の具の飛散跡及びその垂れ方)は相当程度脳内で視覚的にイメージできるから、視覚的思考脳、イメージ脳及び表現脳の活動が結果物全体を因果的に支配していると考え余地もあり、偶然性の要素はどの色がどこに配置されるかとの点にあるにすぎないとも考えられる(また個々のボールの跡に著作物性が認められるかどうかは別問題である。)。このように、個別具体的な評価の性質上、偶然性に関する抽象的な設例から AI 生成物の著作物性を肯定する議論はハードルが高いものであるように思われる。
- 70) 前掲注 67) で示した通り、前掲注 1) 上野・奥邨 157 頁(前田健執筆部分)は、「著作者たるべき者が、最終的な結果物の表現をどこまで意図的にコントロールしていたかは問われない」「著作者と認められるために、結果物の創作的表現を予め意図的にコントロールできることは求められてこなかった」「美術工芸品として美術の著作物に該当するとされている陶芸についても、窯の炎などの偶然性に左右されて作品ができあがる点は、著作物たり得ない根拠とはされてこなかった」と述べる。しかし、この見解は、AI 生成物の著作物性の判断基準において間接正犯類似の道具理論を支持しながら(前掲注 1) 参照)、その理論的支柱である意図的・因果的支配(コントロール)を否定している点で、理論的に難しい問題を孕んでいるように見える。また、前掲注 66) と同様、陶芸の窯変部分に一般的に著作物性を認めることを前提とする議論には難しい面が感じられる。こ

れも個別の事例ごとの規範的評価の問題であって、釉薬の配合や分量、焼き方の加減や時間等を研究して意図的に特定の紋様が顕れるように工夫している場合、偶然性による一定の揺れやずれが生じたとしても、陶芸作品全体に創作意図の因果的支配が認められる余地もあるが、たまたま現れた個々の紋様部分に著作物性は認められないと判断される可能性もある。

- 71) 詳細な長文のプロンプトが言語著作物と評価される余地はあるが、プロンプトの本質は画像の特徴を表す単語又は文節等の羅列であり、欲する画像の特徴を機械に伝えるための符号に類するものであって、それ自体は自己の思想又は感情を表すための最終表現物ではなく、個々の表現もありふれたものと評価されることが多いと思われることから、プロンプト自体の著作物性を肯定するハードルも低くはないように思われる。
- 72) 前掲注 1) 内田同旨
- 73) 本稿は筆者が代表を務める鈴木法律事務所の橋本瑠依氏(法務博士)、佐々木裕香氏(法務博士)及び佐々木亮氏(弁護士)の協力がなければ完成し得なかったものであり、この場を借りて各氏に謝意を表す。

(原稿受領日 2024 年 5 月 31 日)