

## 第1回 著作権法による声の保護

～実演家の権利の正当化根拠を起点に可能性を探る～

弁護士・NY州弁護士・上級VR技術者 関 真也

## I はじめに—検討の背景

## 1. ディープフェイクとは

「ディープフェイク」は、たとえば後掲資料1のように定義される<sup>1)</sup>。

「人をだます」という目的であるとは限らないが、エンタテインメント産業においてもディープフェイクの技術が利用されるようになり、話題を呼んでいる。たとえば、静止画像の生成やアニメーション化、2D写真からの3Dバーチャルオブジェクトの作成等さまざまな目的に利用されるほか、字幕、吹替え等の作業を迅速化することで作品のグローバルな展開を強化する役割も果たすことができる。

俳優・声優・歌手等の外見・声をそっくりに再現する映像・音声等を生成AIで制作することにより、演者のスケジュール調整や予算の問題を解消して世界中の異なる場所で同時に多くの出演の機会を獲得することができ、演者の活躍の場が広がる可能性がある。また、映像作品のメインキャストが不祥事を起こしてしまった場合に、当該作品自体の公開を断念するのではなく、当該キャストの部分だけを生成AI技術によって差し替えることによって当該作品を生かすこともできる。

## 資料1 「ディープフェイク」の定義

「ディープラーニング（深層学習）」と「フェイク（偽物）」を組み合わせた造語。広義には、人工知能（AI）や機械学習によって生成・編集されたメディアやそのための技術のこと。狭義には、人をだます目的で、写真、音声、映像の一部を入れ替えて（別人の顔や声にすり替えるなど）、本物そっくりに合成された偽画像、偽音声、偽映像を指す。

（笹原和俊『ディープフェイクの衝撃 AI技術がもたらす破壊と創造』46頁（PHP研究所、2023年）より）

2. エンタテインメント産業における  
生成AI・ディープフェイクに対する懸念

しかし、ディープフェイクには負の側面もある。たとえば、生身の演者の出演機会や収入を奪うのではないかという懸念である。演者の立場からすれば、本人の才覚・努力等の成果である実演等を無断で学習・利用され、市場や収入を奪われるのは不公平ではないかという捉え方もできよう<sup>2)</sup>。

International Confederation of Societies of Authors and Composers (CISAC) が2024年に公表した調査によれば、生成AIの市場への浸透に伴い、生成AIが存在しなかった場合と比較して、2028年の音楽クリエイターの収益は24%低下し、また、視聴覚作品のクリエイターの収益は21%低下するおそれがあるという<sup>3)</sup>。作品の制作等のために人間に必要とされる専門的なスキルも、自ら作品を生み出すことのみならず、生成AIツールを使いこなしてクリエイティブな視点から作品を監督・監修・制御又は調整するという側面が強調されるようになる可能性がある。

国際労働機関 (ILO; International Labour Organization) のレポートによると、こうした変化は、「労働ダイナミクスに大きな変化をもたらし、伝統的な創造的役

割の需要を減少させる一方で、技術的及びハイブリッドなポジションにおける機会を増加させる可能性があり、これによってリスクリギングが不可欠となり、既存の労働力の適応に課題が生じる可能性がある」という<sup>4)</sup>。

そのほか、本人なら断るであろう演技・歌唱・演奏等を無断でさせられることに対する不快感もあるだろう。実際、米国で声の模倣が争われた裁判例の中には、アーティストとしての芸術的な忠実性を損なうから音楽アーティストはCMに出演すべきではないという哲学を持つ歌手によく似た声を、本人の同意なくラジオCMに利用したことが問題となった事例がある<sup>5)</sup>。

### 3. 対応状況

このような状況のもと、ILOは、クリエイターのAIに関するスキル開発を通じた労働力の移行を管理するための政策が必要であるとともに、「3Cs」(Compensation / Informed Consent / Control) という概念（資料2）に基づき、AIの倫理的かつ公正な利用に関する保護フレームワークを構築する必要があるとしている。海外では、契約を通じた対処が進められるほか<sup>6)7)</sup>、United Voice Artists (UVA) が「AI Guidelines for Voice Work」<sup>8)</sup>を公表するなど、この分野における指針を示す努力も行われている。

また、米国では、州レベル及び連邦レベルにおいて、俳優等の権利保護を強化することを通じてエンタテインメントにおけるAIその他デジタルメディア技術の責任ある利用を確保すること等を目的として、声を含むデジタルレプリカを規制する立法対応が進みつつある<sup>9)</sup>。

たとえば、テネシー州では、2024年3月、AIによる

声の模倣を禁止すること等を内容とする法律が成立した（通称「ELVIS法」）<sup>10)</sup>。

さらにカリフォルニア州では、2024年9月、同意を得ずに死者の声又は肖像のデジタルレプリカを使用した表現的な視聴覚作品又は録音物を制作、頒布又は利用可能にした者に損害賠償義務を課す法律が成立した（知的財産に関する州民法改正 [AB1836]）<sup>11)</sup>。

また、同州では、同月、本人が自ら行ったであろう業務に代えてその個人の声又は肖像のデジタルレプリカを作成及び使用することを認める契約の条項を、当該条項が当該デジタルレプリカの使用目的に関する合理的に具体的な説明を含まないことその他一定の条件を満たす場合に執行不能とすることを内容とする法律が成立した（雇用に関する州労働法改正 [AB2602]）<sup>12)</sup>。

そして連邦レベルでは、米国著作権局が『著作権と人工知能』に関する一連のレポートの第1弾となる「デジタルレプリカ」編を2024年11月に公表したほか<sup>13)</sup>、声や肖像のデジタルレプリカに対処するための連邦法案が提出された（通称「NO FAKE Act of 2025」）<sup>14)</sup>。

わが国においても、音声業界三団体（協同組合日本俳優連合、一般社団法人日本芸能マネージメント事業者協会、一般社団法人日本声優事業者協議会）が生成AIに関する主張（資料3）を発表したほか<sup>15)</sup>、声優有志による『NOMORE無断生成AI』の取組みも行われている<sup>16)</sup>。政府も、現状では法改正による対応という方針が示されているわけではないものの、近時、肖像・声それぞれについて、現行不正競争防止法の考え方から想定し得る適用事例が示された<sup>17)</sup>。

#### 資料2 「3Cs」の概念

- 公正な報酬 (Fair Compensation) : 実演家は、AI生成コンテンツにおける声、似顔絵又は画像の利用に対して公正な報酬を受けるべきである。
- インフォームド・コンセント (Informed Consent) : 契約においてはアーティストが彼らのデジタルレプリカ又は肖像の具体的な利用について知り得ることを保証しなければならず、また、彼らの氏名、イメージ及び肖像のAIアプリケーションにおける利用について明示的な許可を必要とする。これは、彼らの作品がAIモデルの学習又は生成プロセスにおける入力として利用され得るか否か、また、彼らの作品がどのように利用され得るかを特定することによって利用の範囲（例として、彼らの作品に基づくAI生成コンテンツを商業目的又は非商業目的に使用することができるか否か）を決定できることも含まれる。
- コントロール (Controls) : クリエイターは、彼らの作品、アイデンティティ及び創造的な入力が、AIシステムによってどのように利用され、応用され、又は複製されるかを管理することができる必要がある。

### 資料3 生成AIに関する音声業界三団体の主張

- ①生成AI音声を、アニメーション及び外国映画等の吹替では使用しないことを求める
- ②生成AI音声を学習・利用する際は、本人の許諾を得ることを求める
- ③生成AI音声には、AIによる生成物であることの明記を求める

## 4. 本連載の趣旨と想定事例等

もっとも、わが国における声の保護は、「現行法上、パッチワーク的なものにならざるを得ず、音声に対してシームレスな保護を付与する法制度の構築が喫緊の課題」となっていると指摘される<sup>18)</sup>。とりわけ、音声業界三団体の主張においても生成AI音声を使用しないことが求められた「アニメーション及び外国映画等の吹替」のように、作品への出演（いわば彼らの「本業」）に関わる部分における俳優・声優・歌手等の声の保護については、前述した現行不正競争防止法の適用事例においても取り上げられておらず、その保護の可否が明らかにされていないように思われる。

そこで、本連載では、この「本業」の部分に焦点を当て、著作権法（第1回）、パブリシティ権（第2回予定）並びに商標法及び不正競争防止法（第3回予定）それぞれにおける声の保護の可能性及び範囲について検討を加える。

この趣旨を踏まえ、本連載における想定事例と前提条件は、以下のとおりである。

まず、想定するケースは、著名な俳優・声優・歌手等の声を、彼らが実際には出演していない表現作品に登場させ、あたかも実際に演技・歌唱・演奏等をしたかのようなコンテンツを作り出し、販売、配信、上映、放送等により商業的に利用するケースである。

また、上記「表現作品」としては、複数の表現的な要素が複合する表現作品を想定する。典型例として、映画、アニメやゲームは、多数の出演者・キャラクター、セリフ、演出、脚本、カメラワーク、衣装、背景、美術、アイテム、視覚効果、主題歌、BGM、効果音等が複合した表現作品である。

そして、本連載において検討対象とする「声」の要素ないし特徴は、原則として、「大きさ」「高さ」「速さ」「間」及び「音色（音質）」という5つの要素<sup>19)</sup>に限定する。何を言うかという言語ないし発声内容の要素は具体的的事案によって異なり、俳優・声優等はむしろ作品ごとにこれらが異なる表現を求められるところ、ディープフェイクとして出力される音は言語等の要素において学習元の音声とは異なる場合が多いからである。

## 5. 第1回の検討の流れ

本連載第1回においては著作権法上の課題を検討するが、俳優・声優・歌手等を保護する権利として実演家の権利を主な検討対象とし、著作権の検討は省略する。「声は表現をするための手段にすぎず、表現そのものではないということから、特定人の声や合成音声の特徴それ自体を著作物として保護することはできない」と一般に評価されているからである<sup>20)</sup>。

まず、実演家の権利の正当化根拠を起点として、ディープフェイクの事案においてどのような場合に侵害が成立し得るかを総論的に整理する（後述II）。その上で、当該整理を踏まえ、学習段階につき、実演の享受目的とは何か、また、実演家の利益を不当に害することとなるのはいかなる場合かを考察する（後述III）。なお、生成・利用段階については、後述IIの整理が概ね通用すると思われる所以、本稿では割愛する<sup>21)</sup>。

第1回における条文番号は著作権法のものを指す。

## II 実演家の権利の正当化根拠に基づく総論的な検討

### 1. 実演家の権利とは

実演家は、その実演を録音し、又は録画する権利を専有するほか（91条1項）、90条の3以下に規定される権利を有する。ここにいう「録音」とは、「音を物に固定し、又はその固定物を増製すること」をいう（2条1項13号）。

「実演家」とは、「俳優、舞踊家、演奏家、歌手その他実演を行う者及び実演を指揮し、又は演出する者」をいい（2条1項4号）、「実演」とは、「著作物を、演劇的に演じ、舞い、演奏し、歌い、口演し、朗誦し、又はその他の方法により演ずること（これらに類する行為で、著作物を演じないが芸能的な性質を有するものを含む。）」をいう（2条1項3号）。

## 2. 実演家の権利の保護を正当化する根拠

実演家の権利を含む著作隣接権を保護する根拠については、大きく分けて2つの見解がある<sup>22)</sup>。1つは、「著作物の創作に準じた創作活動の奨励を目的に、『著作隣接権』の趣旨を『著作物に準ずる準創作物』の保護に求める見解」（準創作行為保護説）である<sup>23)</sup>。他方は、「著作物等の情報の伝達行為・給付に要する投資保護を目的に、『著作隣接権』の趣旨をその伝達者・給付者の保護に求める見解」（伝達行為保護説ないし投資保護説）である<sup>24)</sup>。両説の相違は、保護法益を創作的要素と解するか否かの点にあるとされる<sup>25)</sup>。すなわち、伝達行為保護説では、創作的な要素がなくとも「実演」として保護され得る<sup>26)</sup>。これに対し、準創作行為保護説においては、創作的な要素がないものは「実演」として保護しない（たとえば、ありふれた動作や朗読（2条1項18号参照）は「実演」に該当しない）と整理される<sup>27)</sup>。

裁判例では、実演家の権利に関して準創作行為保護説に親和的なものが見受けられる。たとえば、「実演、レコードというものについては著作物の創作活動に準じたある種の創作的な活動が行われるものであるところから、そういった著作物の創作活動に準じた創作活動を行った者に著作権に準じた保護を与えることが、その準創作活動を奨励するものであり、そのような著作物に準ずる準創作物の知的価値を正当に評価する、というものであり、また、この著作隣接権の趣旨は、著作権制度を前提として、著作物を公衆に提供する媒体としての実演・録音に知的価値を認め、著作物の解釈者としての実演家と著作物の解釈の伝達者としてのレコード製作者との関係を合理的に調整して、権利関係を定めることにある」と述べた裁判例がある<sup>28)</sup>。

そのほか、[ファッショショ事件] 知財高裁判決は、ファッショショにおけるモデルのポーズ及び動作につき、「その態様もありふれたものにすぎないのであるから、『これらに類する行為で、著作物を演じないが芸能的な性質を有するもの』に該当するものということはできない」と判示している。これを、ありふれた態様であること（言い換えれば、創作的要素がないこと）を理由に「実演」該当性を否定したものと評価するならば、著作隣接権（そのうち少なくとも実演家の権利）の成立に創作的要素を必要とする準創作行為保護説に親和的な立場をとる裁判例として位置付けることもできる<sup>29)</sup>。

次項で述べるとおり、生成AIに関わる著作隣接権の議論においては、これまでのところ、伝達行為保護説をもとにしたものが多いと思われる。本稿では、いずれの

立場が妥当かという議論には立ち入らないが、これまで述べたように準創作行為保護説に親和的と思われる裁判例もあることから、同説も実務上無視することはできない。以下、両説それぞれの見地から検討を進めたい。

## 3. 問題の所在：ディープフェイクによる声の模倣との関わり—伝達行為保護説から

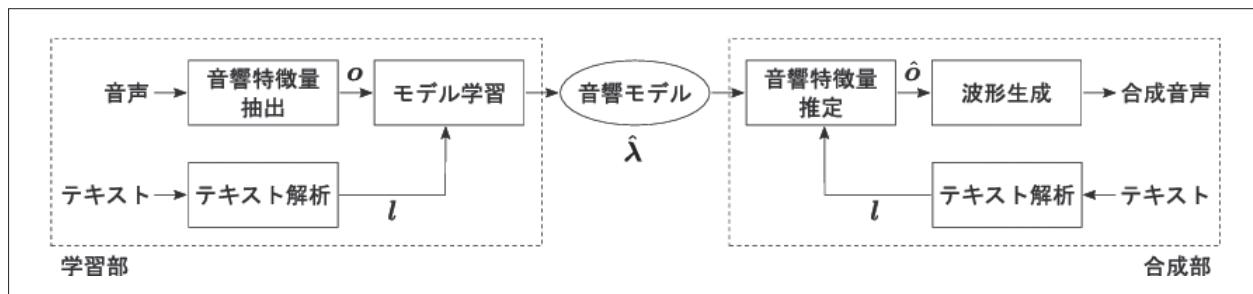
実演家の権利は、著作権とは異なり、表現のレベルではなく個々の「実演」ごとに保護が及ぶ。そして、どのような場合に「その実演」の録音等をした（それゆえに実演家の権利が及ぶ）と認められるかについては、伝達行為保護説の立場から、実演の成立要件として創作性が要求されていないという解釈のもと「元の実演の物理的な連鎖が続いている限り及ぶことになる」としつつ、デジタル技術の進展を踏まえ、元の実演を識別できる程度に再現されていることを条件に、たとえその再現された部分に創作性がなかったとしても実演家の権利の侵害になる（逆に、元の実演が識別できないほどに変容している場合は侵害にならない）旨を述べる見解がある<sup>30)</sup>。

この見解によれば、再現されたのが創作性のある部分であるか否かにかかわらず侵害の成立が認められ得ることになる点では、創作性を要求する準創作行為保護説よりも侵害成立範囲が広いという捉え方ができる。

問題は、ディープフェイクの事例において、いかなる場合に、①元の実演の物理的な連鎖が続いており、かつ、②元の実演を識別できる程度に再現していると認められるかである。

この点につき、「音声化したいテキストの音素ごとの言語特徴量を求めてから、統計モデル辞書と言語特徴量を使用して、テキストに対応する音響特徴量を持つ音声波形を生成する『統計モデル型音声合成方式』を採用している場合、合成された音声は元の音と物理的に同一ではないため、著作隣接権の侵害にはならない」とする見解がある（以下、便宜的に「物理的連鎖断絶説」という）<sup>31)</sup>。

深層学習に基づく統計的音声合成（その概要につき資料4）は、学習部と合成部から構成され、①学習部では、「音声データから音響特徴量系列 $o$ を、テキストから言語情報量系列 $l$ を抽出し、それらの特徴量を用いて音響モデル $\lambda$ を学習」し、②合成部では、「合成用テキストから言語特徴量系列 $l$ を抽出し、対応する音響特徴量系列 $o$ を音響モデルから統計的に予測」した上で、「音響特徴量系列 $o$ に基づき音声波形を合成する」ことによって実現されるようである<sup>32)</sup>。「特徴量」とは対象データ

資料4 統計的音声合成の概要<sup>35)</sup>

の特徴を定量的な数値として表したものであり、「音響特微量としてはスペクトルパラメータと励振源パラメータが用いられる」そうである<sup>33)</sup>。このように、統計的音声合成においては、元の実演に含まれる声その他の音声データは、数値として表される音響特微量という形で抽出される。また、生成される音声は、音響モデルから統計的に予測される音響特微量に基づいて合成される音声波形として出力される。物理的連鎖断絶説は、このプロセスにおいて、元の実演と出力される音声との間の物理的連鎖は断絶されたと評価するものであると思われる<sup>34)</sup>。

これに対し、「元の実演の物理的な連鎖については、直接的なものに限らず、因果関係が途切れていないという程度でも、元の実演が認識可能であれば十分元の実演とその録音物や録画物との対応は確保できる」から、「生成AIが学習済みデータを用いて生成したものから元の実演が認識できる場合には、学習済データを用いて生成されたという事実を以て、元の実演の物理的な連鎖は保たれている」と述べる見解がある（以下、便宜的に「認識可能性説」という）<sup>35)</sup>。この見解はさらに、著作権侵害における依拠の要件に関して具体的にどの表現に依拠したかを要求しないとする裁判例<sup>36)</sup>を参考に、元の実演を具体的に特定せずとも、特定人による実演であるかのように認識され得る音声等が生成されれば、どのルートを辿ったにせよ元の実演の物理的な連鎖が保たれており、実演の録音等と評価して差し支えないとされる<sup>37)</sup>。

従来、①物理的な連鎖に加えて②元の実演の識別可能性が要求されたのは、物理的な連鎖がありさえすればよいという立場を貫徹すると元の実演の原型をとどめないほどに変形している場合でも禁止権が及ぶことになるがそれは妥当ないので、著作隣接権侵害の成立範囲を限定すべきという文脈があった<sup>38)</sup>。これに対し、認識可能性説は、元の実演が認識可能であることをもって、物理的な連鎖を因果関係というレベルで柔軟に捉えることにより、物理的連鎖断絶説に比べると著作隣接権の成立範囲を広く解釈することを志向している点で特徴的であるようと思われる。

たしかに、機械的失業や実演の機会の減少による実演

に対する投資ないしインセンティブの阻害を防止するという伝達行為保護説の趣旨からすれば<sup>40)</sup>、実演家自身や実演を享受する側からは認識できない生成AIの技術的なメカニズムではなく、出力された音声等が元の実演に代替し得るものと彼らから認識されるかどうかが重要になる。技術面がどうであれ、代替可能な実演（類似のもの）であると彼らに認識されれば、実演の機会や意欲が減少するおそれがあるからである。そうすると、認識可能性説のように、元の実演と出力された音声との対応関係の認識に重点を置いて権利が及ぶかを判断した方がよいといえるかもしれない。ただし、元の実演の識別可能性や因果関係の有無等をどのように立証・認定するかが肝心であることはいうまでもない。

#### 4. 準創作行為保護説からの検討

準創作行為保護説によれば、元の実演に表れた準創作的な要素（たとえば、著作物等についての解釈・再構築とその提示<sup>41)</sup>）を観念し、それが共通する範囲内で「その実演」の録音等に該当すると解釈する余地が生じる。

この点、メタバース上のアバターの動きに関してであるが、「実演の保護が、個性の表出にあることも併せて考えれば、アバターがなす実演が『その実演』であるといえるためには、機械的再生であること、アバターがなす実演から、もとの実演の個性の表出を感得でき、当業者を基準にしてその同一性を感得できることが基準になる」とする見解がある（以下、便宜的に「個性表出説」という）<sup>42)</sup>。

もっとも、個性表出説も、元の実演の個性の表出を感得できることに加えて、元の実演の機械的再生であることを要求している。これは、準創作的要素の共通だけで「その実演」を録音したといえるのかという疑問が生じるためだと思われる。したがって、統計的音声合成の場合に、元の実演に含まれる声が、出力される音声において機械的に再生されているかどうかをどのように判断するかという点につき、課題が残る（3. で前述した物理的連鎖断絶説と認識可能性説に関する議論を参照）。

実演の準創作的要素ないし個性に保護理由を求める個性表出説を貫くと、出力される音声の内容が、（固定や増製といった録音その他の実演の利用行為に関わる）機械的ないし技術的なプロセスを経た結果として、元の実演に表れた個性に由来するものといえる限り、機械的再生にあたるといえようか。このように解すると認識可能性説における元の実演の認識可能性や因果関係の問題に近くになると思われるが、統計的音声合成に関してこれらをどのように認定するかが肝心であることもまた同様である。

一方、準創作行為保護説の立場からも、準創作的要素と実演におけるその具体的表現を構成する素材の捉え方によっては、統計的音声合成の出力は元の実演を「録音」（増製）するものではないとする考え方があり得る。この点、実演の（準）創作的要素を、「著作物の全体解釈を前提とするその再現方法における創作的要素」と解した上で、この再現方法自体はアイデアであり保護の対象にならないとし、「実演において、たとえ著作物の再現方法としての発音・発声方法や動作方法が共通であったとしても、実演の具体的表現を構成する素材としての音声、容貌、身体が異なれば、もはや先行実演の権利はおよばないと解すべき」であると述べ、声帯模写等の物真似には実演家の権利が及ばないと説明する見解がある（以下、便宜的に「素材共通説」という）<sup>43)</sup>。この見解によれば、声帯模写（いわゆる「声まね」）は、音声としては本人の実演と同じだが身体（声帯）が異なるため「その実演」の利用に当たらないという整理になると思われるが、これとパラレルに考えると、統計的音声合成により出力される音声は、元の実演を行った実演家とは身体（声帯）等が異なるから、実演家の権利は及ばないことになると思われる。

## 5. 小括

以上の検討によれば、準創作行為保護説と伝達行為保護説のいずれの立場においても、統計的音声合成のケースにおいて、出力される音声が元の実演の「録音」であることを否定する見解も導くこともできるし、「録音」であることを肯定する余地を認める見解を導くこともできる。

もっとも、いずれにせよ、結局は侵害が肯定されるケースは限られるのではないかという課題が残る。なぜなら、たとえ演じる対象の著作物等が同じでも、実演は機会ごとに別個のものであるため、出力された音声から、そのうち特定の実演を識別ないし認識することは困難な場合が多いと思われるからである。

とりわけ準創作行為保護説によると、同一人物が同一著作物等について同様の解釈・再現方法をもって演技等を行った場合、少なくとも1回目は準創作的要素があるから「実演」に該当すると認めやすいが、2回目以降については1回目とは異なる新たな準創作的要素がないため「実演」に該当しないことになるか、又は「実演」に該当すると評価するとしてもそれぞれ別個の実演に表れる解釈・再現方法における準創作的要素を細かく細分化して観念しなければならず、これを元の実演として学習したとして、出力される音声にその準創作的要素が表れていると認定することは困難となるため、やはり保護範囲は狭くなってしまう<sup>44)</sup>。

## III 実演の学習段階

### 1. 論点の整理

統計的音声合成においては、学習対象である元の実演の音声データをコンピュータに入力することになる。この点において、音の固定物の増製、すなわち「録音」（2条1項13号）があると考えられる。これを前提に、生成AIによる声の再現に關し、学習段階で著作権法上問題となるのは、主に、30条の4による権利制限を受けるか否かである（同条は、102条1項により、著作隣接権の目的となっている実演の利用について準用される）。

具体的な論点としては、①実演の何を享受する目的がある場合に同条の準用が否定されるのか（同条本文）、また、②いかなる場合に実演家の利益を不当に害することとなるか（同条ただし書き）の2点が挙げられる。

### 2. 実演の享受とは何か（30条の4本文）

#### 2-1. 問題の所在

30条の4は、「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」ことを権利制限の要件とする<sup>45)</sup>。実演の利用について同条を準用する102条1項は、この場合における読替えの仕方については規定していない。もっとも、同条9項2号は、「第一項において準用する第三十条の四の規定の適用を受けて作成された実演等の複製物を用いて、当該実演等を自ら享受し又は他人に享受させる目的のために、いずれの方法によるかを問わず、当該実演等を利用した者」（強調付加）は、録音等を行ったものとみな

すと規定している。これによれば、102条1項による30条の4の準用に際しては、102条9項2号と平仄を合わせ、「当該実演等」を享受する目的があるか否かが文言上問題になると考えるのが自然であるとされる<sup>46)</sup>。

では、実演の享受とは何をいうのか。言い換えれば、実演の場合、具体的に何を享受させる目的があるときに、30条の4の準用が否定されるのであろうか。

## 2-2. 伝達行為保護説からの検討

前述II.3の整理によれば、伝達行為保護説の立場からは、①元の実演の物理的な連鎖が続いている、かつ、②元の実演を識別できる程度にこれを再現したAI生成物を出力する目的がある場合に、享受目的があると判断するのが基本的な考え方となろう<sup>47)</sup>。準創作行為保護説とは異なり、元の実演の創作的な要素を再現したAI生成物を出力する目的があることは不要となる。

統計的音声合成のケースにおいてどのように考えるかは、物理的連鎖断絶説と認識可能性説とで判断が分かれ得る。

物理的連鎖断絶説によると、統計的音声合成の場合には、そのプロセスにおいて元の実演と出力される音声との間の物理的連鎖は断絶されたと評価されるから、(仮に出力された音声から元の実演を想起し得るとしても)元の実演に含まれる声のデータを学習するに際して享受目的ではないと判断されることになると考えられる。

これに対し、認識可能性説によれば、声を含む元の実演のデータをAIモデルに学習させることにより、当該元の実演を認識できる音声を出力させるという目的があると認められる場合に、享受目的があると判断されることになると考えられる。

しかし、前述II.5でまとめたところによれば、認識可能性説においても、享受目的があると認められるケースは相当限られるかもしれない。別個に行われた同様の声の実演が複数ある場合、そのうちのどの実演を識別ないし認識させる目的であったかを区別するのは困難なことがあると思われるからである（認識可能性説の立場からの立法的対応の示唆につき、前述II.3、注38）参照）。

## 2-3. 創作行為保護説からの検討

前述II.4の整理によれば、準創作行為保護説からの基本的な考え方は、元の実演に表れた準創作的な要素（たとえば、著作物等についての解釈・再構築とその提示）が共通するか否かであり、これが共通すれば「その実演」の「録音」等に該当すると解し得る。伝達行為保護説とは異なり、AI生成物において元の実演を識別できるとしても、その準创作的な要素を再現する目的でな

ければ享受目的ではないと判断されると考えられるため、この点において30条の4による権利制限を受けやすいといえる。

もっとも、創作行為保護説をベースにする場合であっても、個性表出説と素材共通説とでは、やはり享受目的の有無に関する考え方方が異なり得る。

個性表出説によれば、統計的音声合成により出力される音声から元の実演の準創作的な要素ないし個性を感じできるかどうかが基準となるが、加えて機械的再生であることも必要となる（どのような場合に機械的再生があると認められるかについては、前述II.4を参照）。前述II.5でまとめたところによれば、個性表出説においても、前述2-2と同様の理由で、享受目的があると認定するのはそう簡単ではない。

他方、素材共通説において、準創作性を基礎付ける具体的表現の素材として実演家の身体（声帯）を重視するすれば、統計的音声合成により出力される音声は、元の実演を行った実演家の身体という表現の素材が共通しないから、そもそも同一の準創作的な要素を感じできないため享受目的は否定されると考えられる。

## 3. いかなる場合に実演家の利益を不当に害することとなるか（30条の4ただし書き）

### 3-1. 著作権に関する議論を参考に

前述2のとおり、いずれの見解に立ったとしても、統計的音声合成によるディープフェイクの場合には、実演を享受する目的があることを理由に30条の4の準用を否定することはできないか、又はできるとしてもそのケースは限られる。したがって、実演家の権利を根拠として声の保護を求める権利者の立場からは、同条ただし書きが適用されるかどうかが重要なポイントとなる。

この点、いわゆる作風や画風が共通する場合における著作権に関する議論では、「著作権法が保護する利益でないアイデア等が類似するにとどまるものが大量に生成されることにより、特定のクリエイター又は著作物に対する需要が、AI生成物によって代替されてしまうような事態が生じることは想定しうるもの、当該生成物が学習元著作物の創作的表現と共に通しない場合には、著作権法上の『著作権者の利益を不当に害することとなる場合』には該当しないと考えられる」とされる<sup>48)</sup>。

本稿が想定する「声」の5つの要素も、それ自体は実演そのものではなく、実演を構成し得る一素材であるという点では、作風等に関する上記議論に類するところがある。とりわけ、準創作行為保護説のうち素材共通説は、

著作物の再現方法自体はアイデアであって保護の対象にはならないとし、実演の準創作的な表現とは区別しているから、実演に関する30条の4ただし書きの議論も作風等に関する上記議論に準じることとなる。すなわち、実演における具体的表現の素材として実演家の身体的要素を重視するすれば、統計的音声合成によるディープフェイクの場合には身体（声帯）という素材を含む準創作的な部分が共通せず、再現方法というアイデアが共通するにすぎないため、実演家の利益を不当に害することとなるとはいえないという結論になりやすいと考えられる。また、身体的要素を重視するかどうかという点で差はあるかもしれないが、個性表出説についても、個性ないし準創作的要素が再現されているか、それともそれに至らない部分のみが共通するにすぎないかが結論を左右することになると考えられる。

他方、伝達行為保護説では、「実演」として保護されるために準創作的要素が必要でない。このため、著作権の場合と異なり、30条の4ただし書き該当性に関して、生成AIやディープフェイクによって準创作的な表現が再現されていることは必須条件ではない。物理的連鎖や元の実演の識別可能性があるといえる範囲内で実演が再現されているかどうかが問題となる。

しかしながら、前述III.2のとおり、出力された音声において、元の実演の個性ないし準創作的要素が再現しているか、又は物理的連鎖等がある範囲内で元の実演が再現されていると評価できるケースは限られると思われる。結果として、ただし書きにより30条の4の準用が否定されることは少ないとされる。

### 3-2. 学習行為との因果関係に着目した不当な利益侵害の評価

以上の整理をもとに実演家の立場から見ると、ディープフェイクによる声の模倣に関しては、元の実演の個性ないし準創作的要素が再現されているかどうか（準創作行為保護説）、又は物理的連鎖や元の実演の識別可能性があるかどうか（伝達行為保護説）にかかわらず、俳優・声優・歌手等の声に代えてディープフェイクが「表現作品」（前述I.4）に使用されることによる（潜在的）市場の衝突がもたらす出演機会の減少ないし喪失<sup>49)</sup>という彼らへの不利益をもって、「実演家の利益を不当に害することとなる場合」に該当するという立論を試みる必要がある<sup>50)</sup>。

30条の4において実演家の権利の侵害の成否が問題とされるのは、あくまで学習行為である。元の実演を「情報解析の用に供する」ために録音することが実演家の権利の侵害となるかが問題なのであり、AIが生成した音

声の出力や利用がそれ自体で権利侵害であるかどうかは、本来の問題ではない。同条ただし書きは、情報解析の用に供するために録音される「当該実演」（「当該著作物」からの読替え）の種類及び用途や、「当該利用」すなわち情報解析に必要と認められる限度における当該実演のAIモデルへの入力（録音）の態様に照らして、実演家の利益を不当に害することとなるかどうかを問うものなのである。

そうだとすれば、実演家の利益を不当に害することとなるか否かは、統計的音声合成により出力される音声が、それ自体、実演家の権利を侵害するものである（すなわち、元の実演の個性ないし準創作的要素が再現されている／物理的連鎖や元の実演の識別可能性がある）かどうかではなく、元の実演の学習行為と因果関係のある範囲内で実演家の利益が不当に害されているかどうかによって判断すべきである。

出力される音声それ自体が実演家の権利を侵害するものである（ことが多い）という事実は、実演家の利益が不当に害されていると判断するための重要な材料の一つになり得るが、そう判断するための不可欠の事情とはいえない。そもそも、前述III.2のとおり、出力された音声に元の実演の個性ないし準創作的要素が再現されているか又は物理的連鎖等があるということは、元の実演を「享受」できることを意味するが、ただし書き該当性は、かかる「享受」の目的がないために本文に該当することを前提として問題となる。すなわち、ただし書きは、実演家の利益を不当に害することとなる場合であるか否かを判断するに際し、「享受」の目的がないこと、すなわち、統計的音声合成によって出力された音声において元の実演の個性ないし準創作的要素の再現又は物理的連鎖等がないケースを当然に想定しているといえる<sup>51)</sup>。それにもかかわらず、出力された音声にこれらがあることを、実演家の利益を不当に害することとなると認めるための不可欠の事情とする解釈はできないというべきである。

このように解することができるとすれば、特定の俳優・声優・歌手等の「声」を含む実演をAIに学習させたことと因果関係のある状況において、その特定の実演家の「表現作品」への出演機会がディープフェイクに代替され、（潜在的）市場の衝突や彼らの実演の利用に対する阻害が生じた場合には、出力・利用される音声において、学習された彼らの実演の準創作的要素の再現又は物理的連鎖や元の実演の識別可能性があるか否かにかかわらず（著作権のケースに置き換えていえば、アイデアが共通するにすぎないか、それとも創作的表現まで共通するかにかかわらず）、彼らの利益を不当に害するものとして30条の4の準用が否定される余地がある。

さらに試論を進めれば、著作権の場合に比して、実演家の権利の場合、とりわけ伝達行為保護説から考えるとには、アイデアと創作的表現の区別は関係なく、実演に対する投資ないしインセンティブが阻害されないようにするため、機械的失業や実演の機会の減少から実演家を守るという観点から、ただし書きにおける「実演家の利益」への影響（市場の衝突や実演の利用の阻害等）を広く捉えることも正当化されるかもしれない。そこで課題となるのは、ディープフェイクに係る学習行為によって実演家の市場が奪われるという事実と因果関係の立証であろう<sup>52)</sup>。

30条の4ただし書きに関しては、本稿で検討した点のほかにもさまざまな課題がある<sup>53)</sup>。また、学習済みモデルにはさまざまな利用方法があり、著作者や実演家の利益を不当に害することとならない利用方法も多々あり得るのに、学習段階を広く規制することであらゆる利用方法を制限することは妥当でないという懸念にも留意する必要がある<sup>54)</sup>。

## IV おわりに

本連載第1回では、実演家の権利の正当化根拠を起点として、実演家、A I開発・利用者等のさまざまな立場からの立論の可能性を探った。次回以降も、現行法及び裁判例のもと、生成A I・ディープフェイク時代のエンタテインメント産業における声の保護の可能性を模索する。

- 1) 笹原和俊『ディープフェイクの衝撃 A I技術がもたらす破壊と創造』46頁（PHP研究所、2023年）。
- 2) この点は、法的にはさまざまな捉え方が可能である。本連載で検討するように、声の保護に関して各知的財産法がどのような利益をカバーするかにもよるが、たとえば、時間・労力・費用がかかる技能の修得を通じた成果に対するフリーライドを無制限に許すことは、俳優・声優・歌手等の活動が行われなくなり、エンタテインメントに関わる産業と文化の発展を阻害するおそれがあるとし、一般不法行為の成立を検討する余地があるかもしれない（東京高判令和6年6月19日（LLI/DB判例番号L07920477）〔バンドスコア〕参照）。また、独占禁止法その他の競争法の観点も見過ごせない。連邦取引委員会（FTC）のLina Khan委員長（当時）は、「誰かが作成したコンテンツや情報がスクレイピングされ、その後、それらと競合したり市場から排除したりしてビジネスを奪う方

法で利用された場合、一部のケースでは、それは不公正な競争方法に該当する可能性がある」と述べたという（THE HILL, *Generative AI and the media and culture industry* (Feb. 27, 2025), available at <https://thehill.com/policy/technology/4682461-ftc-chair-ai-models-could-violate-antitrust-laws/>）。

- 3) CISAC, *Global economic study shows human creators' future at risk from generative AI* (Dec. 2024), available at <https://www.cisac.org/Newsroom/news-releases/global-economic-study-shows-human-creators-future-risk-generative-ai>.
- 4) ILO, *Research Brief: Generative AI and the media and culture industry* (2025), available at <https://www.ilo.org/publications/generative-ai-and-media-and-culture-industry>.
- 5) Waits v. Frito-Lay, Inc., 978 F. 2 d 1093 (9th Cir. 1992).
- 6) SAG-AFTRA, *TV/Theatrical Contracts 2023*, available at [https://www.sagaftra.org/sites/default/files/sa\\_documents/TV-Theatrical\\_23\\_Summary\\_Agreement\\_Final.pdf](https://www.sagaftra.org/sites/default/files/sa_documents/TV-Theatrical_23_Summary_Agreement_Final.pdf).
- 7) SAG-AFTRAは、実演家がそのデジタル・ヴォイス・レプリカをデジタルオーディオ広告において利用するためのライセンスを与えることができるようにするオンラインプラットフォームを提供するNarrativと合意しており、これにおいて「3 Cs」を重要な要素に挙げている（<https://www.sagaftra.org/sag-aftra-x-narrativ>）。
- 8) United Voice Artists, *AI Guidelines for Voice Work* (2024), available at <https://unitedvoiceartists.com/wp-content/uploads/2024/05/UVA-AI-GUIDLINES.pdf>.
- 9) 北村弥生「【アメリカ】カリフォルニア州におけるディープフェイクを規制する法律」外国の立法No.302-1, 2頁（2025年）参照。
- 10) Ensuring Likeness, Voice, and Image Security Act of 2024, available at <https://publications.tnsosfiles.com/acts/113/pub/pc0588.pdf>.
- 11) California Assembly Bill 1836, available at <https://legiscan.com/CA/text/AB1836/id/3021237>.
- 12) California Assembly Bill 2602, available at <https://legiscan.com/CA/text/AB2602/id/3021235>.
- 13) U.S. Copyright Office, *Copyright and Artificial Intelligence, Part 1: Digital Replicas* (July 2024), available at <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf>.
- 14) Nurture Originals, Foster Art, and Keep Entertainment Safe Act of 2025, available at <https://www.blackburn.senate.gov/services/files/BBD 6 E069-2ED5-473A-B81A-4815134B876F>.
- 15) 「生成A Iに関する音声業界三団体の主張」(<https://www.nippairen.com/wp-content/uploads/2024/11/25>

- d2c1a91dc0alc23acc1f17ee068af4.pdf)。
- 16) 『NOMORE無断生成A I』のウェブサイト (<https://nomore-mudan.com/>)。
  - 17) 経済産業省経済産業政策局知的財産政策室「肖像と声のパブリシティ価値に係る現行の不正競争防止法における考え方の整理について」(2025年)。
  - 18) 安藤和宏「音声の法的保護に関する一考察」三村量一古稀『切り拓く一知財法の未来』701頁、708頁（日本評論社、2024年）。
  - 19) Gudrun GRÄWE「日本文化における『声』」立命館言語文化研究29巻3号155頁、162頁（2018年）（鴻上尚史『发声と身体のレッスン』（筑摩書房、2012年）を参照）。
  - 20) 今村哲也「声の知的財産法による保護に関する一考察」特許研究79号8頁、12頁（2025年）。
  - 21) 生成A I・ディープフェイクと実演家の権利の問題を取り扱う主な文献として、本稿で挙げるものほか、柿沼太一「生成A Iと声優・歌手の実演」自由と正義2024年6月号25頁（2024年）、唐津真美「講演録 メタバース・生成A Iの時代における実演家の権利と法的課題」コピライト761号2頁（2024年）、古場和美「岐路に立つ著作隣接権制度—生成A Iと実演の関係を中心にして—」26頁（第11回著作権・著作隣接権論文集、著作権情報センター、2024年）、田邊幸太郎「生成A I時代における『声』の保護に関する検討」三村量一古稀『切り拓く知財法の未来』745頁（日本評論社、2024）等参照。
  - 22) 本山雅弘『著作隣接権の理論』218頁、266頁以下（成文堂、2021年）。同「著作隣接権の保護理由と実演における創作的要素」公益社団法人日本芸能実演家団体協議会実演家著作隣接権センター（CPRA）編『実演家概論』29頁以下（勁草書房、2013年）、上野達弘「実演と隣接権制度」論究ジュリ26号12頁（2018年）も参照。
  - 23) 前掲注22) のほか、加戸守行『著作権法逐条講義〔七訂新版〕』627頁（著作権情報センター、2021年）、吉田大輔「著作隣接権制度の可能性」コピライト428号2頁、6頁以下（1996年）も参照。
  - 24) 前掲注22) のほか、田村善之『著作権法概説 第2版』518～519頁（有斐閣、2001年）、安藤和宏「アメリカにおけるミュージック・サンプリング訴訟に関する一考察（2・完）」知的財産法政策学研究23号231頁、279頁以下（2009年）、橋谷俊「判批（2・完）（知財高判平成26年8月28日〔ファッショナリ〕）」知的財産法政策学研究48号329頁、381頁以下（2016年）参照。
  - 25) 本山雅弘「著作隣接権の趣旨はいかに解されるべきか」論究ジュリ26号4頁（2018年）。
  - 26) 張睿暎「生成A Iと著作者及び実演家の権利」獨協法学122号147頁、162頁（2023年）、佐藤豊「生成A Iによる実演の学習、実演類似のものの生成及び生成結果の利用に対する規律の一考察①」コピライト760号26頁、29頁（2024年）。
  - 27) 本山『著作隣接権の理論』前掲注22) 360頁参照。
  - 28) 東京高判平成14年10月17日裁判所ウェブサイト（平成11年(仮)第3239号）。
  - 29) 知財高判平成26年8月28日判時2238号91頁〔ファッショナリ〕本山雅弘新・判例解説Watch17号288頁（2015年）参照。ただし、橋谷・前掲注24) 399頁も参照。
  - 30) 安藤・前掲注24) 282頁。田村・前掲注24) 522～523頁も参照。
  - 31) 安藤・前掲注18) 704頁。
  - 32) 橋本佳＝高木信二「深層学習に基づく統計的音声合成」日本音響学会誌73巻1号55頁～56頁（2017年）参照。
  - 33) 橋本＝高木・前掲注32) 56頁。
  - 34) 張・前掲注26) 175頁は、「A I音声合成物は、歌手が直接歌った実演そのものではないので、実演家の著作隣接権の侵害であるということはでき」と説明する。
  - 35) 〈画像出典〉橋本＝高木・前掲注32) 56頁より。
  - 36) 佐藤・前掲注26) 30頁。
  - 37) 東京地判昭和51年5月26日無体集8巻1号219頁〔サザエさん〕。
  - 38) ただし、これについては立法的な対応（生成されたものが特定人による実演であると認識される場合に、元の実演の具体的な特定を要しないことを明確にする推定規定の新設等）の必要性も示唆されている。佐藤豊「生成A Iによる実演の学習、実演類似のものの生成及び生成結果の利用に対する規律の一考察②」コピライト761号44頁、45頁（2024年）。
  - 39) 物理的な連鎖に加えて元の実演の識別可能性が求められる理由を含め、安藤・前掲注24) 277～279頁及び282頁。なお、田村・前掲注24) 523頁は、録音等の過程で変更が加えられても、「実演の機会の減少をもたらしうる限り権利が及ぶ、換言すれば、元の実演を識別しうる限度で録音、録画の範囲に含まれる」と述べておられる。このように、元の実演の識別可能性は、実演家の権利が及ぶ範囲を一定程度確保するという側面もある。
  - 40) 安藤・前掲注24) 278～279頁及び282頁及び田村前掲注24) 513頁におけるレコードや実演の保護の正当化理由及び範囲に関する議論を参照。
  - 41) 関真也「メタバースにおけるモーションデータの保護と実演家の権利～アバターを通じた『実演』の捉え方とプログラムの位置づけ～」国際著作権法研究2号34頁、40頁（2022年）。本山「著作隣接権の保護理由と実演における創作的要素」前掲注22) 44頁以下、大江修子「講演録：実演家の権利と外延を考える」コピライト701号12頁以下（2019年）も参照。
  - 42) 粟原祐介「メタバースを中心とするバーチャルリアリティにおける著作権法の『実演』に関する一考察—『その実演』の意義を中心に」情報通信政策研究6巻2号15頁、32頁（2023年）。安藤・前掲注24) 283頁も、「準創作的行為を行っていることを実演家に対する権利付与の正当化根拠にする説では、実演に準創作的表現が表れているか否かが保護範囲のメルクマールとなる」と述べておられる（ただし、

- 「準創作的表現が表れている実演とは一体何かという難題をクリアしなければならない」とも指摘されている。)。
- 43) 本山『著作隣接権の理論』前掲注22) 366頁。
- 44) そもそも、学習データが開示されない限り、学習された元の実演がどれであるかを特定することすら困難であろう。
- 45) 非享受目的としては、機械学習を含む「情報解析（括弧内略）の用に供する場合」がその典型例として規定されているが（30条の4第2号）、情報解析の目的がある場合であっても、享受目的が併存する場合には、同条による権利制限の対象にならないと解されている（文化審議会著作権分科会法制度小委員会「A Iと著作権に関する考え方について」10頁、19頁以下（令和6年3月15日））。これに対し、「2号の「情報解析……の用に供する場合」の著作物利用（A Iへの入力行為）は、常に非享受利用に該当する」と理解するものとして、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『A Iと著作権』20頁（勁草書房、2024年）。同じく上野=奥邨編著『A Iと著作権』213頁〔座談会・谷川和幸発言〕も参照。
- 46) 今村・前掲注20) 14頁等参照。これによると、関・前掲注41) 40頁のように、単純に30条の4の「著作物」を「実演」と読み替え、「当該実演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」とするか否かを問題にするのは誤りとなる。本稿では、改めて、実演等の享受とは何をいうのかを、実演家の権利の保護理由を起点にして検討を試みた。なお、「実演等」の定義については102条2項参照。
- 47) 享受目的があると認められるためには、実際には、出力する目的だけではなく、出力した上でそのコンテンツの種類に応じた本来的な利用に供する目的があることも必要であると思われる。たとえば、音声を出力したのであればそれを聞くこと又は聞かせること、画像や映像を出力したのであればそれを見ること又は見せること等である。本稿では、便宜のためこの点は割愛して検討を進める。
- 48) 「A Iと著作権に関する考え方について」前掲注45) 23頁。
- 49) 「A Iと著作権に関する考え方について」前掲注45) 23頁参照。
- 50) 著作権に関し、作風・画風と表現の区別という観点から詳しく述べられるものとして、愛知・前掲注45) 32頁以下参照。
- 51) 「A Iと著作権に関する考え方について」前掲注45) 23頁参照。
- 52) 関連して、米国におけるフェアユースの議論が注目される（U.S. Copyright Office, *Copyright and Artificial Intelligence, Part 3: Generative AI Training, pre-publication version* (May 2025), available at <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>）。
- 53) 上野達弘「A Iと著作権に関する考え方について」法の

支配215号87頁、91~96頁（2024年）、田村善之「生成A Iをめぐる著作権法の課題」知的財産法政策学研究70号13頁（2025年）、奥邨弘司「生成AIと著作権～『考え方について』に関する3つの論点」知的財産法政策学研究70号103頁（2025年）等参照。

54) 上野=奥邨編著『A Iと著作権』前掲注45) 223頁〔座談会・谷川和幸発言〕参照。

#### \*せき・まさや

弁護士・NY州弁護士（関真也法律事務所）。上級VR技術者。

漫画、アニメ、映画、ゲーム、TV、音楽等のコンテンツやファンションのほか、XR（VR/AR/MR）、メタバース、VTuber/アバター、生成A I、eSports、バーチャルファンション等のテクノロジーに関わる領域の知財・法務に従事する。一般社団法人XRコンソーシアム監事、日本商標協会理事・法制度研究部会長、ファンションビジネス学会ファンションロー研究部会長、一般社団法人日本知財学会コンテンツ・マネジメント分科会幹事等を務めるほか、関係する官公庁の委員等の経験も持つ。東海大学総合社会科学研究所客員講師その他大学講師等を歴任。『XR・メタバースの知財法務』（単著、中央経済社、2022年）、『ビジネスのためのメタバース入門—メタバース・リアル・オンラインの選択と法実務—』（共編著、商事法務、2023年）、『ファンションロー〔第2版〕』（共著、勁草書房、2023年）、『バーチャル空間のビジネスと知財法務の教科書』（編集代表、日本法令、2024年）をはじめ、エンタメ、ファンション、メタバース等に関する執筆・講演等多数。