

## 人工知能コンテンツ

——それはビジネスと知的財産権をどう変革するか——

福 井 健 策\*

**抄 録** 人工知能（AI）によるコンテンツやサービスの自動生成は既にビジネスとして大きく広がっている。それらは、生成スピードや膨大性のほか生成経路の複雑化などの特徴を持ち、表現やビジネスのあり方に多大な影響を与える可能性がある。その知的財産的保護は、①学習用データ、②アルゴリズム、③学習ノウハウ、④学習済みモデル、⑤AI生成コンテンツのそれぞれについて、著作権・特許権・営業秘密・契約による保護など新たな問題を提示しており、世界に先駆けて内閣府知財本部などでの議論が始まっている。それは常に、イノベーションのための投資の保護と、一方で成果物への自由なアクセスを守ることによるイノベーションの促進の、最適バランスを持って語られるべき問題である。

### 目 次

1. はじめに
2. AIコンテンツの広がり
3. AIコンテンツの特徴と社会影響
4. AIの知財問題
  4. 1 学習用データ
  4. 2 AI本体（アルゴリズム）
  4. 3 学習ノウハウ（「調教法」）
  4. 4 学習済みモデル
  4. 5 生成されるコンテンツ
5. おわりに

### 1. はじめに

お定まりの将来予測からはじめよう。オクスフォードの2人の教授が、今後10～20年の間に人間が機械に奪われる職を列挙したというあれである<sup>1)</sup>。失職の可能性が高い職業の上位には、データ入力者、電話オペレーター、銀行の融資担当者、タクシー運転手、料理人（!）などが並んだ。最後の料理人は、なにせ国が英国だから、「お前らの国のコックはな！」という世界中のツッコミを浴びたとか浴びないとか。まあ

しかし、ファストフードなどでの調理も含めれば、確かに既に半機械化されているし無理な予測でもないだろう。他方、失職しない可能性が高いと言われた職の上位には、医師・小中の教師・弁護士（←本当か?）などと共に、いわゆるクリエイター・アーティスト系の仕事が軒並み並んだ。「それはそうだ。創作や発明は最も人間らしい営みであって、人工知能やロボットに簡単に奪われてたまるか」と思う方も多だろう。確かにそうかも、しれない。

しかし、スティーブ・ジョブズの意見は違う。ジョブズによれば創作とは「既存の情報の結び付け」に過ぎない。そうであれば人工知能（AI）などはお手の物だろう、とも思える。実際、対象物の特徴を自ら発見し識別・再現などできる「ディープラーニング」の発達で、AIのおこなえる作業領域は急速に広がっている。「判別ビジネス」の広がりはずごい。顔認証・音声認証・文字識別にはじまり、運転中のドライバーの居眠り判定、人や農作物の健康状態判定からボル

\* 日本弁護士 ニューヨーク州弁護士、日本大学芸術学部客員教授 Kensaku FUKUI

本文の複製、転載、改変、再配布を禁止します。

ノ・バイオレンスの自動判別フィルタリングまで、ちょっと検索しただけで幾らでも出て来る。医師や弁護士に関する前述の予測にしても、(ありがたいが実際には) 診断や調査は既にAI的な半自動システムがなくては成り立たない。

## 2. AIコンテンツの広がり

AIによる自動生成されるコンテンツも、既にビジネスとして大きく広がっている。しばしば挙げられるのは「小説の自動執筆」だが、例えば自動翻訳や、半自動撮影の写真を大量活用したグーグル「ストリートビュー」のようなサービスは、既に完全に商用化され広く流布している。

音楽の自動生成でも、BGMの分野は既に実用化済みだ。図1は人気の高い「Jukedeck」だが、誰でもアクセスしてジャンル・曲調・長さなどを指定すると、30秒ほどでBGM音源を自動生成してくれる。



図1 自動BGM生成サイト「Jukedeck」

曲は実際に聞いてみて頂きたいが、同じパラメータを指定してもその都度ユニークな曲が生成され、少なくとも映像、イベントやショップのBGMとしては十分に使えると筆者は感じた。なにせ安い。個人ないし従業員10名以下の企業ならば、出典を明示するだけで無料でダウンロードし自由に使える「ロイヤルティ・フリー」である。既に世界でも活用例は多いだろう。特

徴は、そのすさまじい物量である。1曲が30秒ならば、仮に同時に1曲しか作曲できなくても1年間稼働を続ければ実に100万曲以上作曲できる。JASRACが管理する世界中のプロの楽曲総数の約3分の1だ。これまでとは創作のスピード・物量が桁違いなのだ。

あるいは、画像の加工も活用が広がるジャンルだ。図2は、写真の自動加工サービスである。



図2 「DeepDream」上のAnonymaViktor氏作品 (オリジナルはカラー)

使われているのはGoogleの公開するAI「DeepDream」で、ある画像（ここでは炎）を与えてその特徴をAIに学ばせる。すると、その後で与えた任意の写真（ここではキツネ）を、自動でその作風に変換して「炎のキツネ」を描いてくれるのだ。この間、ものの数十秒であり、サービス自体はやはり無料で提供されている。同じことは、動画でも可能だ。

楽しい？ そう楽しいのだ。AI創作物というと、少し以前には「機械に人間並みの創作が出来るか」という点に関心が集まった。しかし、一流作家クラスの創造をAIが独力で行うまでには、（できるとしても）さすがにもう少し時間がかかるだろう。そのかわり、このように我々が使いこなせるツールとして、我々自身がコンテンツを生み出す助力の役割なら既に十分果たせる状態にある。しかも、マーケット的にはこうした参加型・体験型のコンテンツの方が、ビジネスとしては広がってさえている。

それは時に、逆に侵害を大量に生み出すツールとしても機能する。本稿執筆中に世間を騒がせたのは、「キュレーションメディアの大量閉鎖問題」である。きっかけとなったのはDeNAの人気サイトWELQやMERYだ。ネット上など社会の様々な情報を集めて（キュレーションして）作られた記事が集まる、「まとめサイト」の一種と言え、一時はグーグルなどの各種検索で上位を占める隆盛ぶりだった。

しかし、そこに集まる大量の記事には、時に「肩こりは幽霊が原因の場合もある」といった極めて質の低いものや、ネット上の他の記事のコピペ・寄せ集めに過ぎないものが少なくないと指摘され、事業者側が謝罪、閉鎖に追い込まれたというものだ。しかも、リクルートなど他社のサイト・記事の閉鎖も次々と続く事態となった。

背景に、前述のコンテンツの大量流通があった。AIに限らず、ITネットワーク革命の特徴

は「コンテンツ流通量の圧倒的増大」である。世界75億人が発信者となって、膨大なテキスト・画像・映像が世界を駆け巡っている。数が増えれば増えるほど、その単価は市場原理によって下がる。これが、出版やCDなどいわゆるパッケージ系コンテンツ産業の構造不況の主因のひとつで、「フリー化」と言われる。

価格が下がれば、ますますコンテンツ制作に多額のコストはかけられない。キュレーションメディアでは、だいたい1記事あたり1,000円～2,000円で、学生などのバイトに発注されたという。1文字単価が1円を下回る極端なディスカウントで、プロのクリエイターが受けられる仕事では到底ない。当然、粗製乱造の内容の怪しい記事は増えることになり、コピペも横行した。

そして、その際にしばしば利用されたとされるのが、「リライトツール」や「ボット」と言われるソフトである。既存のテキストなどを読み込ませると、自動で単語などを置き換えて「一見違う文章を作る」ツールだ。図3はそうしたリライトツールの例だ。広告では「人工知能の活用で1秒で2,000文字をリライトする」と謳っている。著作権のクリアされた文章ならばどうリライトしても自由だが、単にネット上から人の文章を漁って来てリライトツールにかけたライターも多かったとされる。



図3 AIによるリライトツールと名乗る広告例

このリライトツール、筆者も試しに使ってみたが、リライト後の文章はとうてい著作権侵害



を免れるなどというレベルには達していなかった。単なる完全コピーではグーグルなどの自動判別で「類似コンテンツ」としてランキング順位が下がるため、それを防ぐためにリライトツールが多用された節もあり、どうもそもそも「法的にセーフなレベルを目指す」という観点自体がなかったようにも思える。仮に、これも「AI」と呼ぶならば<sup>2)</sup>、いわばAIの助けも借りて侵害記事が大量生産されていたと言え、(問題化するまでは) そうしたキュレーションメディアに検索上位独占という目的を達せさせていたことになる。

### 3. AIコンテンツの特徴と社会影響

AIの社会影響は極めて広範に及ぶと考えられ、詳しくは筆者も加わる総務省のAIネットワーク検討会議やその後継であるAIネットワーク推進会議の各種報告などを参照されたい<sup>3)</sup>。ここでは、AIコンテンツやAIを活用したコンテンツ流通に注目して、その特徴や社会影響を簡単にまとめてみよう。

筆者がリクルートAI研究所所長の石山洸氏と共著した論文<sup>4)</sup>では、AIによる創作の特徴を前述のような「自動化」「大量化」のほか、創作過程のどこまでがどのアルゴリズムの貢献によるものか、またどこまでが人間によるものか外部からは判別しがたくなる「複雑化・合体化」、コンテンツが場合に応じて刻々と変化するため、どこまでをひとつの作品ととらえるべきか流動化する「動的化」などと分類している。

そうした特徴の結果、これまでの人間の営みを超えた新たな表現や新たなビジネスが生まれる可能性があることは、極めて重要な影響だろう。現在のコンテンツ過剰の状況は文化産業やプロ・クリエイターという存在を根底からつき崩す可能性が十分にあるが、それもかなりの部分はAIをはじめとするITネットワーク革命の産物である。もうひとつ。AIの判定は類似文

章や類似画像の発見にも大いに活用されている。小保方晴子氏の論文が盗用として問題になった際にも、きっかけは論文の自動比較技術による民間の指摘だった。画像や音楽のマッチング技術も進んでいる。つまり、AIを活用した盗作が可能であることと同様に、AIを活用した盗用・コピーの発見も容易ということだ(リライトはそれを逃れるための技術であるとも言える。まさに技術同士のイタチごっこだ)。

このように、AIはコンテンツの生成・流通だけでなく、それに対する著作権の行使(エンフォースメント)も大きく変える可能性を秘めている。

これらは、裏返せば負の社会影響ともなる。大量のコンテンツ生成は、前述のような価格破壊を招く。AIでそれに拍車がかかれば、プロ・クリエイター達の大量失職は十分あり得る事態だ。既に、カメラマンの世界でそれはかなり現実のものとなっていようし、BGMサイトも作曲ビジネスの価格破壊につながりかねない。何より、ライター達の手取り額が大幅に低下したことはキュレーションメディア問題が物語っている。それらは、ある部分ではライブイベントなど他のコンテンツビジネスへのシフトで補えるだろう。また、技術の進展による不可避の変化であり、新たな創作やビジネスの誕生で補完されることもあり得よう。しかし、我々は同時にプロ・クリエイターが(今まで以上に)食べられない時代の社会影響、それが芸術や教育・研究やジャーナリズムに与える影響を考える必要があるように思う。

加えて、レベルの低いAIやなりすましAIによる、単なる著作権侵害も少なくない。前述の通り、現在程度のリライトツールをもAIにカウントするならば、そのリライトなるものの成果物は、高確率で元の文章の表現の特徴を十分に残しており、よって著作権侵害にあたる可能性が高い。本稿執筆中に話題になったものとし

ては、Everfilter事件があった。これは、ディープラーニングAIを駆使して、写真などを1秒でアニメ風に変換するという触れ込みのアプリだったが、すぐに背景画像などが新海誠監督のアニメ作品の背景と酷似しているという指摘を受け、制作側が謝罪し配信を一時停止する騒ぎに発展した。画風を学んで既存の画像を変換するなどという高度なレベルには達しておらず、単に元のアニメ画像を少し手直しして貼り付けるレベルの作業が行われていたようであり、関係者によってはそんなレベルでAIと呼ばれては誤解が広がるとして、「ニセAI」などと呼ぶケースもある。

#### 4. AIの知財問題

以上の社会的影響を踏まえて、AIの著作権をはじめ知財問題を検討してみよう。ここでは、知財本部の議論に即して、図4のような過程を念頭におきながら議論を進めたい。ただし、この図は完全なるAIによるコンテンツ・サービスの自動生成（computer generated works）を想定しているが、前述の通り、現実には完全な自

動生成だけでなく、AIが人間の活動をサポートする半自動生成的なサービスも同時に隆盛しつつあり、ビジネス的にはこちらが先に全盛期を迎えそうだ。もっとも、機械が人間の創作や作品発信を助ける（machine supported works）という意味ならば、印刷・カメラ・音響・映像・放送・デジタル化・ネットワークと、既に我々は100年以上の長きにわたってその恩恵には浴している。ここでの議論の新しさは、ほとんど人間の手を離れて機械がコンテンツやサービスを自動生成する点にあるので、あくまでそれを中心に論じつつ、機械のサポートによりコンテンツ・サービスの生成や発信が飛躍的な影響を受ける部分にも、適宜触れながら進めることにしよう。

#### 4. 1 学習用データ

##### (1) 単独のデータ

まずはこの図4の一番左、データの知財問題を考えよう。AIは多くの場合、大量の文章・画像・音楽その他各種のデータを学習して（食べて）、そして成長を遂げる。この文章・画像・

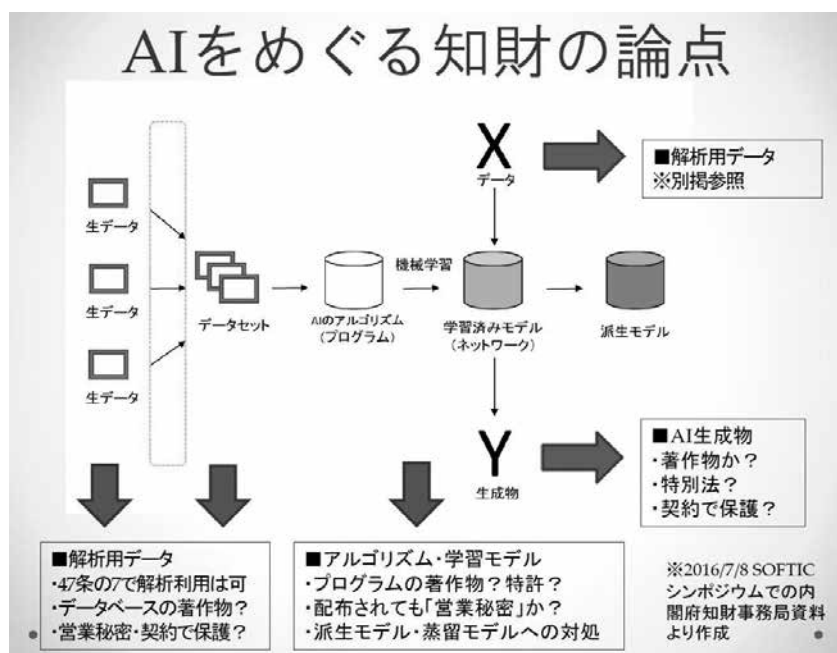


図4 AIをめぐる知財の論点

音楽などは多くのケースでは著作物である。ソーシャルメディア上のつぶやきなども、多くはこれにあたる。よってそれらが保護期間中であれば、原則として他人は勝手に複製・改変などをおこなうことは出来ない。特に業務的・組織的利用の場合はそうだ（著作物としての保護）。

しかし、AIが学習対象とするデータの中には、人々の閲覧履歴・購入履歴・移動履歴や各種のパーソナル・データ、更には気象情報などのノンパーソナル・データといった、著作物にあたらぬ多くの情報が含まれている。これらは個人情報保護法で無断の収集や第三者提供が禁じられるケースもある（個人情報としての保護）。無論、個人を識別できない情報などはその対象でもない。

いずれの場合も、データが秘密として管理などされていれば、その無断での取得や利用は不正競争防止法などによって禁じられる（営業秘密としての保護）。いずれにも該当しなければ、原則として収集も利用も自由だ。

## (2) 集積されたビッグデータ

しかし、こうしたデータは少数では通常必ずしも役立たない。大量に必要なケースが多く、そうなると大量のデータを収集し、管理し、場合によっては系統づけてデータベース化しているものの保護はどうなるか、が次に浮上する。この点は、個別のデータが著作物であろうがそうでなかろうが、データの選択と配列に独創性があればその総体は「編集著作物」として保護される。また、データベース化して検索可能にしている場合、その選択と構成に独創性があれば「データベースの著作物」として保護される。

よって、いずれの場合も、その独創性のあるかたまりを無断でコピーなどすれば、個別の要素に対する侵害の有無とは別に、全体を収集・構築した事業者の編集著作物やデータベース著作物に対する侵害が成立するだろう。かくして、

多くのリソースを投じてビッグデータを蓄積した事業者は、ビッグデータの無断利用から守られることになる。

ただし、ここには「創作性」という留保がつく。これは通常は多くの選択肢があり得るなかから、あえてひとつの選択肢を選ぶことを言う。選択肢がごく限定されている場合、その結果は「必然」であるので創作性は認められにくい。例えば、あるメッセージを伝えるための最も簡潔な表現は、通常は著作権では守られない。そこに選択の余地はほとんどないからであり、見方を変えれば、選択の余地がないような表現を誰かに長期間独占させてしまっただけは危ないからだ。ところが、最近のビッグデータの収集は機械的、あるいは網羅的である。そして機械的・網羅的であればあるほど、収集するデータに選択の幅はなかったことになり、編集著作物性やデータベース著作物性は否定されやすい。

さて、現代においてビッグデータの持ち手とは誰か。あるいは通信事業者や交通機関や通販業者だろうが、最大の保有者は恐らくグーグル、アマゾン、フェイスブックといった巨大プラットフォームだ。つまり以上は、彼らのビッグデータはどこまで独占され得るか、という議論でもある。

では著作権の保護のほか、集積されたビッグデータを独占する手段はないか。前述した、個人情報としての保護は無論あり得るし、また営業秘密としての保護もあり得る。加えて、アーキテクチャや契約による保護がある。例えば、前述のプラットフォームなど多くの企業が保有するビッグデータの場合、公開されているものを除けばそのデータを誰かが持ち出すことは技術的に難しい。つまりアーキテクチャで保護されている。誰かが無断で鍵を破って侵入すれば、もちろん原則として違法となり処罰対象だ。

また、公開されていて一見コピー可能な大量の情報（地図データ、検索結果など）について



も、しばしば利用規約による用途のしぼりがある。例えば、我々はほとんど全員がグーグルのアカウント保持者だろう。つまり、彼らの利用規約に同意をクリックした者ばかりだ。利用規約では、グーグルのコンテンツの複製（あるいは大量ダウンロード）、再配布等はしばしば禁止されている<sup>5)</sup>。無論、そうした事業利用を行う者は多いが、つまりはお目こぼしされているだけだ。ビッグデータ解析などグーグルと直接ライバル関係にある者が大量利用を始めれば、当然、契約違反という主張を受けることになるだろう。最悪はアカウント削除と規約上の罰則に直面することになる。

これは以下のAIの他のファクターの保護の議論にも通底するが、何らかの情報（知的財産）の法的保護を考える時、そこでは単に「権利があるかないか」の二分法ではなく、複数の手段の組み合わせが重要になる。

一方の端には、例えば特別法を作ってビッグデータの無断利用を禁止してしまうとか、その総体を著作物と認めて著作権で独占させるといった、法的権利がある。あるいは個人情報保護法や不正競争防止法も、こちらに入る。法制度による保護だ。そのメリットは、特段相手の合意や権利者側の投資がなくても一律で保護が与えられることで、デメリットは、硬直的で必ずしも最新の社会の動きにすぐに対応できないことだ。変化があまりに急激な情報社会においては、この歩みの遅さは致命傷になりつつある。また、特に日本だけの特別法の場合など、海外では通用しないのもデメリットだろう。例えばビッグデータを保護する特別法を日本でだけ作っても、海外の事業者はコピーし放題だったら、海外でのAI開発に利用され、逆に日本の開発が遅れをとるといった事態になりかねない。

他方の端には、契約やアーキテクチャによる保護がある。前述したように技術的にコピーを困難にしたり、利用規約で商用利用を禁ずる方

法だ。メリットは、柔軟であり、たとえ相手が海外の事業者でも制約なく保護を効かせられることだろう。デメリットは、アーキテクチャはコストがかかる上に破られるかもしれない、また契約は合意した相手しか拘束できない点にある。

よって、目的に応じた両者の使い分けや、両者の協調領域が重要になる。例えばアーキテクチャは技術で破られるかもしれないが、そうしたプロテクション破り行為は、不正アクセス禁止法やコンピュータ関連犯罪の刑法規程で禁じられている。つまり、アーキテクチャによる柔軟な保護を法規制がバックアップしている形だ。

### (3) 利活用促進とのバランス策

さて、以上は主にビッグデータの事業者による囲い込みをどう法的に担保されるかという視点から述べた。しかし同時に、ビッグデータは囲い込むだけが社会善とは限らない。それはAIの「食事」であり、広く共有利用された方がAIに限らず技術開発は活発化しやすいし、囲い込みが強すぎれば国際的な開発競争で後れをとりかねないからだ。

この利活用の促進の方策として、第一に、著作権法の例外規定がある。著作権法には47条の7の制限規定というものがあり、大量の著作物をコンピュータで解析利用することは認められている。これは文字通りビッグデータ解析を指すので、AIによるビッグデータ学習は実は認められている。ただし、最初からデータ解析用に構築されたデータベースは、誰かが無断で利用することは出来ない（同条ただし書）。そのため、現実には本条で解析できないビッグデータも少なくないことになる。

この条文には更に論点があって、自らビッグデータ解析をおこなうために著作物を複製することは確かに認められるのだが、そうして集められた大量の著作物を誰かがビッグデータ解析できるように公衆に向けて提供すること（ネッ

ト公開など)は出来ないように読める(49条1項5号)。

確かに、誰かが著作物を自由に収集してデータベース化し、公衆に向けて公開して良いとなればこれはアーカイブである。それを書籍や音楽、映像について自由におこなわれては目的限定の配信ビジネスと同然ではないか、という懸念があるのはわかる。とはいえ、まず禁じられるのは「解析以外の目的のためのビッグデータの利用」なので、例えばふたつの会社が提携関係を結び、A社が収集してデータベース化した著作物をB社が解析できるようにB社にのみ提供する、つまり特定者間での共有は必ずしも禁止されないと考えて良いだろう。

更には、今後の制度論としてはビッグデータ解析に目的を限定した形であれば、複製した他人の著作物の公衆への提供も広く著作権法の例外規定で認める改正も、考えられて良い。(2017年2月24日に文化審議会 法制・基本問題小委員会が採択した報告書では、こうした二者間の協業や、現行法より広いビッグデータ分析サービスの公衆への提供を認める立法提案がされており、ゆくえが注目される<sup>6)</sup>。)

第二に、契約やアーキテクチャによる過度な独占や規制へのセーフガードとして、あまりに一方的な利用規約が独占禁止法違反とされたり、消費者契約法によって無効とされることも考えられる。逆の意味での法制度と契約の協調と言えるだろう。

第三に、解析用データベースのインフラ整備も進められるべきである。そこでは、利用規約によって解析の用途を限定したり、解析結果が商用利用された場合には権利者側への一定の補償金の後払いを事業者に義務づけるなど、工夫もされて良いだろう。

## 4. 2 AI本体 (アルゴリズム)

次にこうしたデータを学ぶべきAIそのもの、

すなわちアルゴリズムはどう法的に保護されるか。ここでは特にニューラルネットワーク的なものを想定しよう。

これは言うまでもなく、コンピュータ・プログラムとして著作物にあたる場合が多数だろう。一般に著作権法で「アルゴリズム」というと単純な解法を指す用語であって、むしろ(単純過ぎて)著作物にあたらないものの例として挙げられる(10条3項3号)。しかし、ここでいうアルゴリズムはそれよりはるかに複雑な構造を持ったものが大半で、それらは著作物にあたることは疑いない。

また、プログラムは同時に特許としての保護も受け得る。もっとも出願と登録が要件となり、国ごとでの保護となるので、より世界的な保護を指向すれば著作権がベターとなろう。同様に言うまでもなく、それが秘密として管理されていれば営業秘密としての保護も受けよう。

これらも、利活用促進の側面も重要であることは言うまでもない。しばしば、開発者であるプラットフォーム企業などはこのAIをオープンソース・ソフトウェアとして公開し、人々による自由な改良・二次的な創作を認める。著名なものとしてはグーグルが無償公開する「TensorFlow」があるだろう。他の事業者や研究者達は営利目的を含め、これを自由に活用してAIを作成できる。つまり個別のAIを囲い込むよりも、むしろそのAIを開発に活用する開発者達のコミュニティ(=プラットフォーム)になる戦略、だろう。

## 4. 3 学習ノウハウ (「調教法」)

このAIに学習用データを与える訳だが、ここでも学習用データの与え方にノウハウがあるとされ、しばしば「調教法」と俗称される。前述の「炎のキツネ」でいえば、ちょうどよい炎の写真を見つけ出し学習させるからこそ、使える「炎ボット」が生まれるという訳だ。



この調教法はその名の通りノウハウであり、つまりアイデアと思える。著作権法はアイデアを保護しないので、このノウハウは直接には守られない。つまり、炎の写真を与えようというアイデアは誰でも模倣自由である。どの写真を与えたかを秘匿しておけば、つまり学習用データが何であったかを秘密として管理するならば、そのノウハウは営業秘密として保護される。

なお、最近のディープラーニングはAIが自ら大量の対象データを学びその共通の特徴点を学ぶとされるので、今後はこの調教法ノウハウの出番は徐々に減るのかも、しれない。一般に、もはやAIが何を学び、どうしてそのような結果を出力するようになったか、開発者にも完全な把握は困難であるとされる。

## 4. 4 学習済みモデル

### (1) 著作権での保護

次いで、この学習によって生まれた学習済みモデルが挙げられる。つまり「炎ポット」であり、通常は特定のAIと、パラメータ（重みづけ）として表現された関数から成る、とされる。ざっくりと言えば、一定のデータを有するプログラムである。

これは著作物だろうか。AI自体がプログラムの著作物なのだから、学習済みモデルも当然著作物ではあろう。ただ、ここでの問題は、元のAIを離れて、「データ+AI」という新しい著作物が生まれたと考えるかであり、その著作権はどこまで及ぶか、である。逆に言えば、誰かがAIそのものは盗まずに、そのパラメータの部分、データの部分だけを借用する行為は著作権侵害か、である。

これは現在の通説的解釈からはあたらなそう。著作権法は、それ自体が創作的表現とは言えないデータは保護しないのだ。また、例えば全体は著作物にあたるような実録小説や論文で

も、そこから生の歴史的事実やデータそのものを抽出して借りる行為は著作権侵害にはあたらない典型的な行為であり、この点では最高裁判例も確定し堅牢である。よって、学習済みモデルを誰かが適法な手段で学び、そこからパラメータ部分だけを抜き取って利用するとしても、著作権侵害には当たらない。この点は、「あるAIとあるパラメータが結びつく」というアイデアだにとらえても、結果は同じだろう。著作権法はアイデアを保護していない。

別な視点で述べよう。グーグルがTensorFlowでアルゴリズムを無償公開する。人々がそれを学習させて多くの新たな学習済みモデルを作るとする。しかし、それは二次的著作物ではない。新たな著作物は生まれておらず、単にグーグルの著作物である元のアルゴリズムに新たなデータが付着しただけである。そう考えると、成果物は全てグーグルの著作物となりそうだ。

### (2) 特許での保護

では特許はどうか。こちらはアイデアも保護する。また、特許法が対象とする「プログラム等」にはプログラムのほか、「電子計算機による処理の用に供する情報であってプログラムに準ずるもの」も含まれる(2条4項)。よって、あるAIとあるパラメータが結びついている全体を特許として登録し、その保護をAIとパラメータの結びつき自体にも及ぼすなら、パラメータの無断利用を防止できるかもしれない。もっとも、特許は登録が条件で、かつ国ごとの保護である。しかも、その内容を出願書類に記述できることが登録の条件となる。果たして実効的な保護が可能かは疑問もあろう。

更に、著作権にせよ特許にせよ、別な課題もある。例えば、ある学習済みモデルに第三者が更に学習させると、以前とは異なる結果を出力する「派生モデル」が生まれる。これは、同じAIが異なるパラメータを持った状態であり、

恐らく従前の学習済みモデルのコピーとさえ言いくらい。つまり、「元のAI自体の保護」は当然どこまでもついて回るだろうが、「元の学習済みモデルの保護」はついて来にくい。(それ以前に、どのモデルを元にしたかも立証困難だろう。)

あるいは、元の学習済みモデルをコピーしなくても、特定の学習済みモデルに「与えられた元データ」と「結果としての生成物」を大量に把握しておけば、そこから帰納的に同じような機能を有する学習済みモデルは作成できる、とされる。いわゆる「蒸留」といわれる手法だが、これなどは元の学習済みモデルの内容はアクセスすらされておらず、いずれにしても著作権や特許でこれを制約することは一層難しいだろう。

### (3) 営業秘密、契約

では営業秘密としての保護はどうか。これはどこまでいっても秘密として管理されている限りは可能だ。ただ問題は、この学習済みモデルが(これはAI自体でも同様だが)何らかの商品として販売され、ユーザーの手元に移った際、それでも営業秘密として保護されるかである。相手に渡したものを事業者は秘密として管理していると言えるのか? 現に管理できていないからこそ、中身を見られてしまい、模倣されてしまうのではないか? という疑問だろう。

しかし、秘密としての管理とは何も物理的な所在の問題ではなかろう。不正競争防止法自体が、例えばA社がB社と秘密保持契約を結んで自社のノウハウを開示し、そのノウハウをB社が他社に売りとばしたような場合を典型的な営業秘密の侵害、不正競争行為と規定している(2条1項7号)。そもそも、秘密が自社の金庫の中にだけある状態など想定していないのである。よって、例えば販売の際の規約で転用を禁じたり、あるいは暗号などのセキュリティで中身を見られない状態を作っていれば、それは秘

密管理性を満たすはずであり、学習済みモデルは営業秘密として管理するのが、もっとも簡明であるように思う。

この点からすれば、利用規約のような契約自体での保護も当然有効だろう。何であれ、学習済みモデルを公開したり他社に提供する際に契約合意を行い、蒸留や派生モデル作成を含めて転用を制限したり逆に一定の範囲で許してその利用ルールを定めることも事実上可能であり、契約こそが柔軟で簡易な手段であるようにも思える。

## 4. 5 生成されるコンテンツ

さて、このAIが生成する大量のコンテンツはどうだろうか。コンテンツといっても、例えば機械学習したAIに目的地までの最短ルートを表示して貰ったり、AIが将棋の必勝法を編み出した場合、生成物は「ルート」や「必勝法」であり、それらはアイデアである。たとえ人間が考案したとしても著作物にあたらないし、AIが生成しても同様だろう。恐らくこうした種類の多くの生成物は、いかなる知的財産権でも保護されない。

では、もっと作品的なもの、例えばBGMや自動生成の記事はどうだろう。あるいはAIが出力した「炎のキツネ」は、元のキツネ写真に対する二次的著作物となって、そこには新たな著作権が生まれるのだろうか。

この点は、実は1970年代には既に米国CONTUというレポートで論じられており、その後ユネスコや、日本の著作権審議会(当時)でも1993年に議論されている<sup>7)</sup>。結論は本質的には一緒だ。machine supported works, つまりコンピュータをまさにツールとして生身の人間が創作したものは著作物だろうが、computer generated works (CGW), つまり人間は直接的にはボタンを押すだけの完全自動生成のコンテンツは、創作とはいえ著作権の保護は受けない、

である。

唯一の例外は1988年の英国改正著作権法で、そこでは完全なるCGWを著作物と認めている<sup>8)</sup>。が、今もって他国による追従例は目立たないし、筆者の加わった内閣府知財本部「次世代知財システム検討委員会」でも、現行法上は著作権の保護を受けないだろうという点で委員間に異論は少なかった<sup>9)</sup>。

こうした生成物は通常は秘密として管理もされていないだろうから、営業秘密としての保護も難しく、多くは特許対象にもならないだろう。

では、日本も著作権法の解釈を改めたり、特別法を作るなどしてAI生成物に保護を与えるべきか。筆者は現時点ではややネガティブだ。何らかの情報に新たな知財の保護を与える行為は、概ね次のふたつの要素の相関関係で正当化される。つまり、①その情報を独占させることのメリットが情報囲い込みによるデメリットを上回り、かつ、②よりデメリットの少ない代替手段がない場合、である。

では、AI生成物を無断コピーから守るメリットとは何か。AI本体の場合には比較的わかりやすかった。投資された開発コストの保護である。言い方を変えれば、開発に投資するのはその成果物であるAIを利用して収益をあげ、投資回収をはかれるからであり、そのためにはAIを無断コピーからある程度守る必要がある、という訳だ。

では、AI生成物を無断コピーから守らないと投下資本の回収はできないのか。確かにそうした場面もあろう。例えばBGM生成サイトの中には、そのBGMを商用利用する際に使用料を徴収するところがある。前述のJukedeckもそうだ。人々が無断でBGMを転用流用できるならばお金を払うユーザーはいなくなり、開発や運用を続けられなくなる恐れがあろう。(これは多くのコンテンツのビジネスモデルと同様だ。)

ただ、デメリットも同時に気にかかる。なに

せAIコンテンツはその生成ペースが常軌を逸している。年間に100万曲である。少しの投資で近い将来には年間1億曲だって不可能ではなくなるだろう。そのような莫大な著作物が生み出され一事業者がその独占利用権を世界的に約100年もの長期にわたって掌握するという事態を、現行著作権制度は全く想定していない。理論上は、それらのAIコンテンツと似た新たなコンテンツの生成は(そのAIコンテンツに依拠したものであれば)著作権侵害となり、時には刑事罰すら伴う結果を招く。そもそもコンテンツ量の爆発による既存コンテンツ産業やプロ・クリエイターへの影響もさることながら、そのことの表現萎縮効果など、デメリットが読めない。

また、AIコンテンツが仮に著作物にあたるとする場合、その莫大な著作権を握る最有力候補は誰だろうか。言うまでもなく、AI開発で独占先行する巨大プラットフォームである。仮に知的財産権を与えとしても、それも少数事業者に寡占されそうである。

では翻って、こうしたプラットフォームはAI生成物への知的財産権を求めているだろうか。意見を聞けば、「貰えるものなら欲しいですね」と仰るかもしれないが、本当に必要か、筆者はやや懐疑的である。なぜか。そんなところでは稼いでいないからである。いわばビジネスモデルが違う。例えばグーグルであれば、コンテンツの個別課金には恐らく大きな関心はあるまい。莫大なコンテンツが集まる場を設け、そこに人々を引きつけ長期間滞留して貰う。そうすることで更に大量のビッグデータが集まり、彼らに向けてターゲティング広告を打つなど、はるかに効率の良いビジネスが可能になる。

どうも、AI生成物について法的独占を論ずる必要は現時点では薄いように思える。しかも、日本だけが新制度を作ったところで国際的な保護への道は遠い。それよりも、アーキテクチャ



や契約によって過剰な流用・転用を禁じ、その利用方法についてコントロールを効かせる方が、効率が良いようにも思えるところだ。

## 5. おわりに

以上、AIの各要素、ビッグデータ、AI自体、調教法、学習済みモデル、そしてAI生成物のそれぞれについて現行法や新たな法制度での保護の可能性と、利活用促進とのバランス策について私見を述べた。前者については、ビッグデータ、AI自体については現行著作権法でも一定の保護が及び、更に学習済みモデルまでは営業秘密としての保護が有望に思えた。他方、AI生成物は無保護になる可能性が高かった。しかし筆者はいずれの要素についても新法などでの保護にはやや消極的で、むしろオープン・クローズ戦略を生かせる契約の一層の洗練こそが実効性のある手段であるように思っている。

また、AIをめぐってはこのほか、あるいはより喫緊で重要な論点として、AIによる新たな権利侵害、またAIを活用した権利のエンフォースメントの変化、更にはAIを活用した新たな「知財トロール」の出現などの論点もあるが、紙面が尽きたのでまたの機会に譲りたい。

いずれにしても、イノベーションのための投資の保護と、一方で成果物への自由なアクセスを守ることによるイノベーションの最適化という知的財産制度の課題は、ここでも変わらない。ただし、圧倒的に変わったものがひとつある。それはスピードである。これまでのような法的制度の議論スピードでは全てがあまりに遅すぎて、我々は勃興する新たな技術とビジネスモデルの後を、2周遅れで単に付いて行っているだけに思える。知財制度は（更にいえば情報ルー

ルにおける政府の役割は）そのスピード故に早晩、「情報社会においてお荷物以外の何らの貢献もしていない」実態を明らかにする恐れがある。それに代わって浮上しつつある情報のルールメーカーこそ、プラットフォームと彼らのアルゴリズムではないのか。知財立国・情報立国を本気で考えるならば、抜本的に変わるべきは、従来の政策議論のあり方かもしれない。

## 注 記

- 1) カール・フレイほか「雇用の未来」(小林雅一「AIの衝撃」(講談社現代新書電子版・2015年)466頁より)
- 2) 狭義の「人工知能」の定義に合致するレベルでは到底ないが、「人工知能」の概念は広く、その裾野には含まれるだろう。現に、このレベルのものをAIと呼ぶのは(研究者の反発に関わらず)通常用語法である。
- 3) 総務省・AIネットワーク化検討会議報告書2016「AIネットワーク化の影響とリスクー智連社会(WINS(ウインズ))の実現に向けた課題ー」(総務省HP)ほか参照
- 4) 福井健策・石山洗「AIネットワーク化の近未来予測と知的財産権」『年報知的財産法2016-2017』(日本評論社)
- 5) Google マップ / Google Earth 追加利用規約(2015年12月17日更新版)2条:禁止行為など参照
- 6) 文化審議会 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権利制限規定の在り方等に関する報告書」(2017年,文化庁HP)
- 7) 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(1993年,著作権情報センター(CRIC)HP)
- 8) 同法178条
- 9) 同報告書(2016年,知財本部HP)

(原稿受領日 2017年1月26日)