



## AI時代における知的財産権に関する BSA | The Software Allianceからの意見

2023年11月 2日

### 総論

BSA | The Software Alliance<sup>1</sup> (BSA | ザ・ソフトウェア・アライアンス、以下BSA) は「AI時代における知的財産権」に関して、知的財産戦略本部（以下、知財戦略本部）に意見を提出する機会<sup>2</sup>を得られたことに感謝します。

BSAは、世界のソフトウェア産業を代表する主唱者で、長年にわたり著作権の効果的な保護を支持してきました。知財戦略本部も認識されているように<sup>3</sup>、AIはデジタル化・デジタル技術の活用を加速させ、日本全体の生産性向上のみならず、様々な社会的課題の解決にも資する可能性があります。著作権法と同様に、AIはコンテンツ・クリエイターを含む、あらゆる産業分野に恩恵をもたらします。

日本の長期的な経済成長と安定のためにAIはきわめて重要であり、日本政府が法的整備状況を引き続き注視しながら、著作権法および関連する知的財産権法により、AIの継続的発展を促すことを確実にしていくことが大切です。

社会的課題に対応するために社会全体のデジタル化を実現し、イノベーションを育む環境を継続的に醸成していくという政府目標の達成を支えるため、知財戦略本部の検討項目の①及び③について、以下、BSAからの意見を述べます。

<sup>1</sup> BSAの活動にはAdobe, Adobe, Alteryx, Altium, Amazon Web Services, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, Cloudflare, CNC/Mastercam, Dassault, Databricks, DocuSign, Dropbox, Elastic, Graphisoft, IBM, Informatica, Juniper Networks, Kyndryl, MathWorks, Microsoft, Nikon, Okta, Oracle, Palo Alto Networks, Prokon, PTC, Rockwell, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Unity Technologies, Inc., Workday, Zendesk, and Zoom Video Communications, Inc.が加盟企業として参加しています。詳しくはウェブサイト (<http://bsa.or.jp>) をご覧ください。

<sup>2</sup> <https://public-comment.e-gov.go.jp/servlet/Public?CLASSNAME=PCMMSTDETAIL&id=095230820&Mode=0>

<sup>3</sup> [https://www.kantei.go.jp/jp/singi/titeki2/ai\\_kentoukai/gijisidai/siryous.pdf](https://www.kantei.go.jp/jp/singi/titeki2/ai_kentoukai/gijisidai/siryous.pdf)

## ① 生成AIと著作権の関係について、どのように考えるか。

AIは幅広い技術とユースケースを網羅していますが、AIシステムは、AIモデルの訓練に使用できる大量のデータのコンピュータ解析から学習します。このモデルは、その上で、新しいデータを含むクエリ（query）が提示されたときに予測を行います。生成AIモデルの中には、これをさらに一歩進め、テキストや画像、あるいは新たなコードなどの新しいデータを生成するものもあります。どのようなタイプのAIが利用されていても、著作物及び個人データのプライバシーや企業秘密情報を尊重することは、責任あるAI開発の礎となるべきです。

AIの訓練にはさまざまなアプローチがありますが、多くの大規模言語モデル（LLM）は、訓練データからデータ（テキスト、画像、もしくは他のデータ等）を「トークン」に変換し、この「トークン化」されたデータ間の統計的相関関係を見つけます。<sup>4</sup> ソースデータは、合法的にアクセスされたデータセット、オンラインで自由に利用できるコンテンツ、その他のソースから得られる場合もあり、著作権保護の対象である場合も無い場合もありますが、AIモデルの訓練におけるそのようなデータの利用は、特定の作品の表現内容とは無関係です。言語スキルを学習するために書籍が利用され、それがデータベース管理の改善に利用されることがありますが、これらの利用は、著作権法第30条の4<sup>5</sup>（以下、法30条の4）に規定される「著作物に表現された思想又は感情の享受を目的としない利用」に該当します。

多様なAIモデル開発の実践と潜在的ユースケースにおいて、法30条の4は、AI機械学習モデルの訓練における著作物の利用が、単なる事実が高く依存していることに対応した、適切な柔軟性もたらした法理です。著作物の保護、また、公益目的で新しく有用なサービスや著作物を開発・創造する上で要される柔軟性、この両面において、現行の法律、特に法30条の4は、正しい指針となります。AI開発者がAIシステムの訓練のために、一般的に利用可能なコンテンツに合法的にアクセスすることを可能とする、この法制度を日本が維持することを我々は支持します。

生成AIと著作権の関係について回答するにあたり、まず、この項目に関する理解を促すため、AIの活用、AI開発のライフサイクル、AIモデルを訓練するための著作物の利用について説明します。さらに、AIシステムを用いた生成物の著作物性、およびAIシステムを用いて著作権を侵害する作品が生成された際に問われる可能性のある法的責任についての見解を述べます。

### A. 著作権との関係におけるAI

AIの機械学習は、様々な産業や用途で利用するために開発または導入された膨大な数

<sup>4</sup> [https://www.axios.com/newsletters/axios-ai-plus-ff6d6e95-f062-495e-aa45-813f95ceae11.html?utm\\_source=newsletter&utm\\_medium=email&utm\\_campaign=newsletter\\_axioslogin&stream=top](https://www.axios.com/newsletters/axios-ai-plus-ff6d6e95-f062-495e-aa45-813f95ceae11.html?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axioslogin&stream=top)

<sup>5</sup> <https://elaws.e-gov.go.jp/document?lawid=345AC0000000048>

の技術を包含しています。機械学習は、将来のデータ入力に基づいて予測や推奨を行うことができるモデルを開発するために、相関関係、パターン、その他のメタデータを識別するための訓練データのコンピュータ解析に依存します。<sup>6</sup> 生成AIモデルは、これにより新たなテキスト、画像、音声を生成することが可能となります。

## A-1 : AIの活用

機械学習におけるコンピュータ解析から得られる洞察、予測、その他の出力について説明するために、以下に広く認識されている活用例をいくつか挙げます。

- 気象条件、リアルタイムの燃料消費量、航空機の運航データ、近隣の航空交通状況、空港の混雑状況、その他数多くのデータ要素のコンピュータ解析に基づく、自動飛行管理と航空交通管制。<sup>7</sup>
- 腫瘍学、自己免疫疾患、パーキンソン病やアルツハイマー病の分野における早期診断、予防、治療のための化学的また細胞的異常の特定。<sup>8</sup>
- セキュリティ・オペレーションに生成AIを統合し、セキュリティの異常を効果的に特定し対処し、潜在的な脅威を検出することによる緩和。
- 衛星データ、測候所データ、地形情報、各種IoT・センサーデータのコンピュータ解析に基づく気候予測モデリング。<sup>9</sup>

<sup>6</sup> 「バイアスに挑む：AIの信頼性構築に向けたBSAのフレームワーク」（2021年）

<https://ai.bsa.org/wp-content/uploads/2021/07/2021bsaaibias.jp.pdf>

米国国立標準技術研究所（National Institute of Standards and Technology, NIST）の

NIST Risk Management Framework（Jan. 2023）も参照ください。

<https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

（AIシステムを「与えられた目的に対して、実環境または仮想環境に影響を及ぼす予測、推薦、又は意思決定などの出力を生成することができる工学的または機械ベースのシステム」と記述しています。）

<sup>7</sup> -M. Durgut, Artificial Intelligence and Air Traffic Control, Aviationfile.com website（2023年1月）

[https://www.aviationfile.com/artificial-intelligence-and-air-traffic-](https://www.aviationfile.com/artificial-intelligence-and-air-traffic-control/#:~:text=One%20of%20the%20primary%20applications%20is%20to%20help,make%20informed%20decisions%20on%20routing%20and%20scheduling%20flights)

[control/#:~:text=One%20of%20the%20primary%20applications%20is%20to%20help,make%20informed%20decisions%20on%20routing%20and%20scheduling%20flights](https://www.aviationfile.com/artificial-intelligence-and-air-traffic-control/#:~:text=One%20of%20the%20primary%20applications%20is%20to%20help,make%20informed%20decisions%20on%20routing%20and%20scheduling%20flights)

-Degas et al., A Survey on Artificial Intelligence and explainable AI in Air Traffic Management, 12 Applied Sciences 1295（2022年）

<https://www.mdpi.com/2076-3417/12/3/1295>

-Hanneke Weitering, How Artificial Intelligence is Transforming Aviation, Futureflight.aero website（2023年）

<https://www.futureflight.aero/news-article/2023-07-13/beyond-automation-how-artificial-intelligence-transforming-aviation>

<sup>8</sup> -Hunter et al., The Role of Artificial Intelligence in Early Cancer Diagnosis, 14(6) Cancers 1524(2022年)

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8946688/>

-Stafford et al., A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases, 3 NPJ - Digital Medicine 30（2020年）

<https://www.nature.com/articles/s41746-020-0229-3>

-Diogo et al., Early diagnosis of Alzheimer's disease using machine learning, 14 Alzheimers Research and Theory 107（2022年）

<https://alzres.biomedcentral.com/articles/10.1186/s13195-022-01047-y>;

<sup>9</sup> -Schneider et al., *Harnessing AI and computing to advance climate modelling and prediction*, 13 Nature

- 輸送ログ、メーター検針、燃料購入記録、大気汚染追跡、発電所やその他の施設の目視監視などのデータソースのコンピュータ解析に基づく、炭素追跡とその緩和。<sup>10</sup>
- 脆弱な海辺の地域をマッピングするためのコンピュータ解析を実施し、大暴風リスク・マップの作成や、大暴風対応シェルター、学校、保健施設やその他の災害計画と生存のためのインフラ投資計画の案内。<sup>11</sup>
- 予測タイピングやその他のオフィスにおける生産性を上げるソリューション（例えば、タイピストが「sin」という英字を入力した後に「...cerely yours」という英字を提案する「オートコンプリート」機能）<sup>12</sup>、あるいはクリエイターや映画プロデューサーの新作の制作を支える音響効果や特殊効果の作成。<sup>13</sup>
- 上述の活用のように、幅広い領域や応用で、人々がAIと協働することを可能にする会話型インターフェースを生成AIが実現。会話型インターフェースにより、人々がAIシステムのデータや出力を照会したり推論したりすることが可能。

ユースケースは多岐に亘るものの、各々を訓練する要素は、後述するように非常に類似しています。訓練の目的は、著作物の保護されない要素をモデルが学習できるようにすることです。

## A-2 : AI開発のライフサイクル

AIの開発ライフサイクルには、通常、以下の手順が含まれます。

- プロジェクトの構想：まず、AI開発チームはシステムが対処しようとする「問題」を定式化し、システムが予測しようとする構造とターゲット変数をマッピングします。特定のタスクのために訓練されたモデル（狭義のAI活用とも呼ばれる）の例としては、フィットネスアプリがあります。ユーザーの心拍数を解析し、その人が脳卒中や心臓病のリスクにさらされているかどうかを予測する

Climate Change 887 (2023年) <https://www.nature.com/articles/s41558-023-01769-3>

-World Economic Forum, *The role of machine learning in helping to save the planet* (2021)

<https://www.weforum.org/agenda/2021/08/how-is-machine-learning-helping-us-to-create-more-sophisticated-climate-change-models/>

<sup>10</sup> Global Data Alliance, *Cross-Border Data Transfers & Environmental Sustainability* (2023年) <https://globaldataalliance.org/wp-content/uploads/2023/04/04192023gdacbdtsustainability.pdf>

<sup>11</sup> Global Data Alliance, *Cross-Border Data Transfers & Environmental Sustainability* (2023年)

<sup>12</sup> -S. Ashraf, *Desmystifying Autocomplete*, Towards Data Science (2020年)

<https://towardsdatascience.com/index-48563e4c1572;>

-A. Wickramarachchi, *Machine Learning: Word Embedding and Predicting*, Towards Data Science (2020年)

<https://towardsdatascience.com/machine-learning-word-embedding-and-predicting-e603254e4d7b>

<sup>13</sup> D. Nelson, *AI Researchers Design Program To Generate Sound Effects For Movies and Other Media*, Unite AI Website (2022年)

<https://www.unite.ai/ai-researchers-design-program-to-generate-sound-effects-for-movies-and-other-media/>

上での不規則性（つまり、ターゲット変数）を監視するのです。しかし、基盤モデルの場合は、幅広い活用を支えるために開発されます。

- 生データの特定： 第二に、AI開発チームは、後に変換され構造化される「生データ」の関連領域を特定します。データソースは、機械学習AIの活用の可能性と同じくらい多様です。M2M（機械間）データ（衛星通信データ等）や国際貿易統計から、出版物、ブログ投稿、ウェブサイトのコメント、チャットルームのログまで、あらゆるものが含まれます。「生データ」はしばしば「乱雑」であり、以下に概説するように、データを使用可能な形に変換するために多大な作業を要します。特定の活用のための狭義なAIシステムの開発、もしくは、特定のタスクのための基盤モデルのファインチューニングのためのデータは、特定のタスクに関連づいています。しかし、基盤モデルの訓練には、より広範で多様なデータが必要になります。大規模言語モデルを訓練するために必要なデータの規模は膨大であり、この規模の達成にはインターネット上のデータ解析が可能となることが不可欠です。
- データセットの準備： AI開発チームは、機械が意味上のつながりを理解できるようにし、モデルの訓練に使用できるようにするために「生データ」を修正することがあります。この過程で、AI開発チームは必要に応じてデータを修正し、クリーニングし、正常化します。通常、データは「トークン化」によって意味的にも構造的にも変換されます。トークン化とは、テキストやデータの一部を、コンピュータ解析のために、より小さな単位（または「トークン」）に分解することです。解析のために、データの信頼性、品質、適合性を改善するのに追加処理が必要な場合があります。これによりデータセット全体において、欠損値、重複、外れ値、一貫性のないフォーマットなどの品質上の課題に対処することができます。大規模AIモデルは通常、自己教師あり（**self-supervised**）手法で訓練されるため、訓練データにラベル付けをする必要がなく、通常インターネットから収集される膨大な訓練データでモデルを拡大させることが可能です。訓練データは最終的に、AIコーパス全体に広がる一連のトークンとして現れることが多いです。
- モデルの定義： 入力データが適切に処理された後、AI開発チームはシステムの基本アーキテクチャを確立しなければなりません。これには、アルゴリズムが将来予測を行うルール的基础となるパターンや関係を探る際の判断となる訓練データ内の変数（すなわち「特徴」）の特定が含まれます。また、システムを強化するアルゴリズムモデルのタイプ（線形回帰、ロジスティック回帰、ディープニューラルネットワークなど）<sup>14</sup>の選択も含まれます。データの準備が整い、アルゴリズムが選択されると、AI開発チームはシステムを訓練して、将来

<sup>14</sup> 「バイアスに挑む：AIの信頼性構築に向けたBSAのフレームワーク」 infra note 6.



のデータ入力について予測を行うことができる機能的なモデルを開発します。

- モデルの検証、テスト、修正：モデルの訓練が完了したら、AI開発チームはそのモデルが意図したとおりに動作しているかどうかを検証し、システムの出力が予想されるパラメータの範囲内にあり、予期せぬエラーや意図しないバイアスが含まれていないことを実証するためにテストを行わなければなりません。検証とテストの結果によっては、AI開発チームはこれらのリスクを軽減するためにモデルを修正し、改良する必要があるかもしれません。

## B. AIモデルを訓練するための著作物の利用

上述したように、コンピュータ解析は通常、数百万から数十億のトークン化されたデータ要素からなる大規模な訓練データのコーパスに適用されます。モデルの訓練方法によっては、インターネット上でアクセス可能なデータが、コーパスに変換される生データセットの一部として収集されることがあります。

この生データには著作物が含まれる可能性があります。インターネット上で自由に利用できるコンテンツの大半が著作権保護の対象となる可能性があるのは、「独創性」を立証するための閾値が低く設定されており、著作物が登録されていなくとも、作品の創作時に自動的に著作権が発生するためです。しかし、重要なのは、オンライン上のすべての資料が著作権の対象となるわけではないということです。著作権保護は事実、アイデア、数学的概念には及ばないからです。<sup>15</sup>

コンピュータ解析は、著作権法に関与する可能性のある二つの複製を含む場合があります。(1)「訓練データ」のコーパスを作成するために必要な複製、(2) AIモデルを訓練するコンピュータ上の処理に付随する一時的複製です。いずれの場合も、複製は、可視化されず、公衆の利用を可能としないという意味で「中間物」でありませぬ。それどころか、複製は、基礎となる作品コーパスにおいて、著作権で保護されない情報を特定する技術的プロセスに必要となる副産物であります。AIモデルの作成において、相関関係やパターンを伝え、これにより、将来のデータ入力に基づく予測が可能となるのです。このような中間的で非表現的な複製は、著作権が保護しようとする経済的利益と競合せず、これに何ら影響を及ぼしません。

さらに、コンピュータ解析においては、その表現的内容のために著作物が享受されることはありません。むしろ、そのような解析に関係するのは、トークン化されたデータセット全体における確率、相関、傾向、その他のパターンの数学的計算です。このような解析は、データセット全体に分布された数学的パターン（たとえば、他のトー

<sup>15</sup> TRIPS協定の9条(2)は「著作権の保護は、表現されたものに及ぶものとし、思想、手続、運用方法又は数学的概念自体には及んではならない」としています。言い換えれば、著作権の保護は、著作物に含まれるいかなる情報やアイデアも対象とするものではなく、そのような情報やアイデアが表現された独創性のみを保護するというものです。したがって、新しい作品を創作する目的も含め、作品に含まれる情報を誰もが自由に利用することができます。

クンとの関連における特定のトークンの関係性)のみを理解しようとするものです。これらの数学的パターン自体は、著作権法によって保護される表現的コンテンツではありません。

以下に述べるように、確立された法理により、このような複製は一般的に許容されています。

### B-1: AI訓練データのコンピュータ解析はデータ解析である

法30条の4の権利制限規定は、パターン、相関関係、その他のメタデータを識別するために、大規模に収集された情報を解析する目的での著作物の利用を確実にするものです。これにより、将来のデータ入力について予測を行うAIモデルを開発することが可能となります。

例えば、自然言語処理モデル(AIによるタイピング予測モデルなど)を作成しようとするAI開発者が、一般に利用可能なテキストベースの資料から訓練データベースを作成することができるということです。このような場合においては、AI開発者は、著作物の表現享受目的でテキストを複製することはありません。むしろ、何千ものフレーズ、言葉の形、比喩、隠喩、文法パターン、一般的な言語的定式や表現に現れる文字や単語間の相関関係、パターン、関係性等の保護されない情報を抽出するためだけに複製が行われます。文字、単語、言い回し、また、何千、何百万もの文章にわたるそれらの間の数学的なパターンは、著作権の保護対象でははなりません。

同様に、画像生成モデルの場合、モデル開発者は、単語でタグ付けされた非常に大量の画像、例えば、「猫」とタグ付けされた猫の写真や、「犬」とタグ付けされた犬の写真を取り込みます。時間の経過とともに、モデルは特定のパターンが「猫らしさ」の特徴であることを学習し、与えられた画像が猫であるかどうか、画像が漫画であるか写真であるかなどを認識するようになります。繰り返しになりますが、機械は表現的な利用のために猫の特定の画像を複製しているのではなく、猫とは何かを理解するために画像を分析しているのです。表現というより、事実や統計的な相関関係の基本的組み合わせの認識です。

さらに、データセットに適用されるコンピュータ解析の最終的な用途は、元のコンテンツの変換です。予測タイピングソフトウェアのオートコンプリート機能は、ソフトウェアコードという形で新たに構成されたものであり、AI学習コーパス全体とも、そのコーパスの中のどの作品とも一線を画すものです。よく使われるフレーズの語尾を提案できるソフトウェアコードは、基礎となる生データにある著作物からは何段階も離れています。最後に、このような機能は、著作権が保護しようとする著作物と競合することは全くありません。したがって、AI開発者は、著作権の制限規定により、合

法的にアクセスできるテキスト、画像、その他のデータを使ってAI訓練データベースを構築することができるのです。

### C.AI システムによって生成された資料・作品の著作物性

AIシステムを用いて生成された作品の著作物性に関しては、その表現を支えるためにどのような機器や技術が使用されたかに関わらず、人間の創造性が作品にどのように寄与したのかが分析の基準となるべきです。他のソフトウェア・アプリケーションが長い間、アーティストや作家の重要なツールであったように、生成AIは創造性を強化すべきです。生成AIは、例えば、作家による言語処理やビジュアル・アーティストによる写真補正に使われ、オーディオビジュアル作品の特殊効果や録音のための音楽アレンジにも使われています。人間の創造性を高めるために生成AIが利用される場合、結果として生じる作品は著作権によって保護されるべきです。

著作物性に関する問いは、人間の創造性の程度を精査することとなり、AIがあらゆる分野の創造的な作品にツールとして使用されることが増えるにつれ、AIが利用された場合に著作物性を制限するという決定は、AIソリューションの採用を大きく阻むこととなります。

AIシステムの生成物として出現し、人間の創造性の要件を満たす作品は、引き続き著作権保護の対象とすべきです。多くの場合、人間の作者やクリエイターが創造的なビジョンを実現するために使用するツールとしてAIシステムは機能します。例えば、写真家は画像編集の面倒なプロセスを自動化するためにAI対応ツールを使い<sup>16</sup>、建築家はエネルギー効率を高める設計を補強するためにAIを使い<sup>17</sup>、映画制作者はアニメキャラクターの動きをよりリアルに見せるためにAIを使うでしょう<sup>18</sup>。いずれの場合も、人間のユーザーの創造的な寄与があるため、その生成物には著作物性があると容易に結論付けることができます。

利用されるのが生成AIであることで、この分析を変えるべきではありません。もっとも、人間の創造性が全く関与していないことが明らかな極端なケースや、逆に生成AIが創造的表現の一部ではない場合もあるでしょう。

### D. AIシステムを用いて生成された侵害作品に対する潜在的責任

AIシステムを利用して生成された侵害作品に対する法的責任の可能性に関する、BSAの見解は以下となります。

---

<sup>16</sup> <https://theblog.adobe.com/adobes-general-counsel-makes-the-case-for-ai/>

<sup>17</sup> <https://www.autodesk.com/redshift/machine-learning-in-architecture/>

<sup>18</sup> <https://theblog.adobe.com/state-of-ai-in-animation/>



## D-1: AIが生成したコンテンツに起因する場合を含め、現行の著作権法は著作権者を侵害から保護している

著作権者は、その権利が侵害された場合、十分かつ効果的な救済を受けるべきです。この原則は、AIシステムを用いた生成物にも、その他の方法による生成物にも等しく適用されます。このような侵害が発見された場合は常に、アーティストやクリエイターに生じた損害を十分に補償することが重要です。

我々は、侵害の課題に対応する上で、現行の著作権法が適切であると考えます。上述したような大半のAIのユースケースでは、AIシステムの出力が著作権に関与することは全くありません。しかし、他のテクノロジーと同様に、AIも著作権を侵害するような資料や作品を作成するために使用される可能性があります。そのような場合、侵害責任は、侵害しているとされる生成物が「実際の複製」に基づくものであり、著作物と「実質的に類似している」ことが証明されることが前提となります。つまり「通常の合理的な人が二つの作品を区別できない」ような、著作物間の「質的にも量的にも著しい」類似性です。具体的な事実を示す方法が一つしかない又は数通りしかないため、著作権の主張が「薄い」場合は、一般的な法規範がそうであるように、生成物と著作物との「事実上の同一性」を証明する必要があります。生成AIの利用者の多くは、新たなツールを試している小規模事業者であり、法的執行の決定はこの点を考慮すべきです。また、二次的著作物を含むような、その他の場合においても、適宜、責任は生じるでしょう。

原告は、AI関連サービスの提供者に対して侵害訴訟を起こすかもしれません。この場合においても、現行の著作権法が、責任の評価と分担において適切であることが証明されるはずですが、原告が直接侵害の事実を証明した場合、裁判所はサービス提供者がそのユーザーの行為に「寄与」したか、「間接的に」責任を負うべきかどうかを判断することになります。

AIモデルの生成物が、モデルが訓練された作品と実質的に類似している場合、裁判所は、その生成物が侵害にあたるかどうかを案件ごとに判断するための手段を自由に行使できます。我々は、現行法が適切かつ公正な境界を設定しており、現時点では政策変更する理由にはならないと考えています。

## D-2: 非常に大規模なデータセットで訓練されたAIモデルは、侵害的な生成物を出力する可能性が低い

訓練データとAIシステムの生成物を混同しないことは重要ですが、訓練に利用できるデータが多ければ多いほど、（ユーザーの侵害意図がない場合においては）システムが特定の入力の複製物や派生物を生成する可能性が低くなることは特筆すべきことで

す。非常に大規模なデータセットで訓練された、うまく設計されたAIモデルであれば、コンピュータ解析によって特定の著作物と「実質的に同一」はおろか、「実質的に類似」した出力が生成されることは絶対に（又はごく稀な状況に限定してしか）ないはずです。

AIの開発者や導入者の中には、アーティストやクリエイターと関わりを持ち、変化するデジタル環境の中での彼らの作品をどのように支えるかを考え、利用できるプロンプトに制限をかけるなど、AIシステムの悪用を制御するための措置を講じているところもあります。

### ③生成AIに係る知的財産権のリスク回避等の観点から、技術による対応について、どのように考えるか。

まず、AIシステムが著作権を侵害する作品の創作に利用されるリスクについてですが、著作権侵害責任の有無を判断するのに著作権法は十分な柔軟性を有しています。別の言い方をすれば、生成物が侵害にあたるかどうかは、利用される技術とは無関係です。生成物が侵害している場合、侵害複製物の作成にAIが使用されたか否かに関わらず、明らかな責任が発生するはずです。

また、AIシステムの訓練を含め、いかなる目的であれ、オンライン上の著作物へのアクセスを実質的に制限する規制の仕組みを作ろうという呼びかけがあることにも触れておきます。自動化された「クローリング回避」ツールはすでに存在し、AIシステムの訓練を含め、いかなる目的であれ、オンライン上の作品へのアクセスを防ぐことは可能です。効果的でコンセンサスに基づいた技術的な仕組みにたどり着けるよう、クリエイターとAIの開発者および導入者間の自発的な対話をさらに進めることを我々は支持します。

## 結論

BSAは、知財戦略本部に意見する機会を得られたことに感謝します。上述の通り、AIの訓練データセットに含まれるコンテンツの享受を目的としない解析は、著作権法の権利制限規定の対象です。BSAは、AIの訓練プロセスに関するマルチステークホルダーの取り組みや、侵害のリスクを最小化する取り組みを支持します。また、侵害が発生する場面においては、コンテンツ・クリエイターが十分に保護されることをBSAは強く支持します。また、職業上および商業上の利益と競合するような、個人の名前、画像、肖像、音声の不適切なデジタル複製に対して、現行の著作権法がより良く保護できるよう検討する取り組みをBSAは支持します。