



「AIと著作権に関する考え方について（素案）」に対する BSA | The Software Alliance からの意見

2024年2月12日

総論

BSA | The Software Alliance (BSA | ザ・ソフトウェア・アライアンス、以下BSA)¹は、「AIと著作権に関する考え方について（素案）」（以下、素案）に関し、文化庁（以下、貴庁）に意見を提出する機会²を得られたことに感謝します。

BSAは、政府やグローバル市場において、世界のソフトウェア産業を代表する主唱者です。BSAの会員企業は、世界で最もイノベティブな企業であり、人口知能（AI）や機械学習システムを含む、エンタープライズソフトウェアのソリューションを開発し、経済のあらゆる分野で、様々な規模の企業の刷新と成長を支援しています。³エンタープライズソフトウェアの開発者を代表し、BSAは効果的な著作権保護を長年、支持してきました。

人工知能（AI）は、デジタル化とデジタル技術の活用を加速させる可能性を秘めており、日本の生産性向上や社会課題の解決に貢献します。AIはあらゆる分野でイノベーションと創造性を促進しています。例えば、映画の特殊効果、音響、建築計画、自動車のスタイルやデザインなどで、AIはクリエイターに新たなツールを提供し、彼らの技能を向上させています。このテクノロジーが進化し続ける中、イノベーションを奨励し、クリエイターの権利を保護する上で、著作権法が果たす役割を考えることは重要です。⁴

¹ BSAの活動には Adobe, Alteryx, Altium, Amazon Web Services, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, Cloudflare, CNC/Mastercam, Dassault, Databricks, DocuSign, Dropbox, Elastic, Graphisoft, Hubspot, IBM, Informatica, Kyndryl, MathWorks, Microsoft, Nikon, Okta, Oracle, PagerDuty, Palo Alto Networks, Prokon, Rockwell, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, 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。

² <https://public-comment.e-gov.go.jp/servlet/Public?CLASSNAME=PCMMSTDETAIL&id=185001345&Mode=0>

³ 「How Enterprise Software Empowers Businesses in a Data-Driven Economy（エンタープライズソフトウェアがデータ駆動型社会においてビジネスを強化する）」:

<https://www.bsa.org/files/policy-filings/011921bsaenterprisesoftware101.pdf> (英文)

⁴ 「Artificial Intelligence & Copyright Policy: Advancing Technology and Creativity in the 21st Century Economy (AIと著作権政策: 21世紀経済におけるテクノロジーと創造性の促進)」

<https://www.bsa.org/files/policy-filings/01082024bsaaincopyright.pdf> (英文)

日本の長期的な経済成長と安定にとってAIがきわめて重要であることを踏まえ、日本政府がこうした便益を最大化し、日本の著作権法（以下、法）および関連する知的財産法が引き続きAIの発展を促進し、著作権者を侵害から保護することが大切となってきます。

この点、著作権者（クリエイター、実演家）、事業者（AI開発者、AIサービス事業者）、AI利用者を含む様々な関係者からの懸念に答えるべく、現行法の解釈について指針を示そうとする貴庁の現行の取り組みを我々は高く評価しています。貴庁も認識されているように、生成AIに限らず、著作権法の解釈は、本来、個別具体的な事案に応じて司法により行われるべきです。しかし、AIの発展と利用を予測することは、日本におけるAIの導入をさらに加速させるために極めて重要です。AIの開発・導入を含む、社会全体のデジタル化を通して社会課題に応じ、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、政府の目標達成を支えるため、以下、BSAからの意見を述べます。

我々は、素案の多くの部分において、生成AIシステムからの生成物と学習データとしての入力が高に不適切に混同されているのではないかと懸念しております。大規模言語モデル（LLM）やその他の生成AIシステムにおいては、通常、数百万、もしくは、数十億のデータを用いて学習が実施され、著作物の一部がいくつかが含まれることもありますが、検索や複製のためにその表現内容を保存することはありません。その代わりに、コンテンツ（例：生データ）を意味解析や構造解析が可能な構造に変換した後（「トークン化」と呼ばれるプロセス）⁵、学習データは統計的パターンのために解析されるのです。つまり、抽出、もしくは、「学習」されるのは、コンテンツそのものではなく、数百万、数十億の著作物の学習データセット全体の解析による、コンテンツに内在するパターンなのです。

現行の著作権法

[2. 検討の前提として / 3. 生成AIの技術的背景について / 4. 関係者からの様々な懸念の声について]

素案に記されているように⁶（4頁）、現行の著作権法は、電子計算機による情報解析、また、AIモデルの学習のための著作物の利用に関する例外規定（法第30条の4、第47条の4、第47条の5）を含め、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、バランスのとれた指針となっています。

上述したように、AIモデルの学習においては、電子計算機による解析では通常、データをトークンに変換し、学習データセット全体において、他のトークン化されたデータとの統計的相関関係を解析します。データセットには数百万、もしくは、数十億のデータが含まれており、このデータの一部は著作物から学習されることもありますが、その利用は通常、著作物の表現内容とは無関係です。AIモデルの生成物が、モデルが学習した著作物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せん。そのため、現時点で法改正は必要ないと考えます。

⁵ 「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」をご参照ください。
<https://www.bsa.org/files/policy-filings/jp11022023iprightsai.pdf>

⁶ 素案2. 検討の前提として（1）

学習・開発段階

「非享受目的」に該当する場合について

[5. 各論点について / (1) 学習・開発段階 / イ 「情報解析の用に供する場合」と享受目的が併存する場合について]

AI 学習に関する既存の例外に対する過度な制限の回避：著作権法と生成 AI を含む AI の関係を解釈しようとする、貴庁の取り組みを我々は高く評価しています。一方、法第 30 条の 4 に基づき、AI モデルの学習に著作物を利用するために設けられた既存の例外の範囲を過度に狭めたり、制限するような状況にすることは避けることを推奨します。

例えば、素案には「生成・利用段階において、学習された著作物と創作的表現が共通した生成物の生成が著しく頻発するといった事情は、開発・学習段階における享受目的の存在を推認する上での一要素となり得ると考えられる」と記されています(18 頁)。⁷同記載は、著作権侵害の可能性のある生成物の頻度が、学習データの利用目的と直接関連するものと貴庁が捉え、法第 30 条の 4 の権利制限が適用されないという考えを示しています。

しかし、生成・利用段階においては、AI の利用者が、定められた利用範囲内で自由にプロンプトを入力し、生成物を生成することが可能であることを理解することが重要です。通常、利用者と AI モデルとの対話・相互作用は、AI モデルを開発・訓練した主体によって管理・制御されることではないのです。ほとんどの LLM では、ユーザーが特定の生成物を生成するようシステムに指示するプロンプトがなければ、特定の生成物の生成が「頻発」することはありません。

利用者が、通常とは異なるプロンプトを使って、AI モデルを騙し（例：指示する）、学習データやその他の著作物と類似したコンテンツを生成させることがあるかもしれません。AI モデルがそのようなプロンプトを通じて著作物と実質的に類似した生成物を生成した場合、侵害が起こりうる対象は、その指示された生成物です。法第 30 条の 4 およびその他の関連する例外規定に基づき AI 学習に著作物を利用することの有効性に、このような侵害生成物を誘発する目的で利用者が行う意図的なプロンプトエンジニアリングが影響を及ぼすべきではありません。

また、基盤モデルは様々なアプリケーションをサポートしていることを理解することも重要です。データセットの解析を通じて学習が実施された基盤モデルは、ある主体によって開発され、その基盤モデルを利用するアプリケーションを開発する別の主体に提供されるかもしれません。モデルの最終的な利用によって生成される生成物の性質をもって、学習データとの関連性の推認をすべきではありません。

開発・学習段階に関する法の評価において、生成・利用段階との関連性を貴庁が一般化しないことを我々は求めます。そのような著作権法の解釈は、AI の開発や学習を行う事業者の予見可能性や法的安定性を大きく損なわせ、その結果、日本における AI モデルの開発や学習を阻害し、AI の開発や導入への投資を減少させることとなります。

⁷素案 5. (1) イ (イ) 非享受目的と享受目的が併存する場合について

検索拡張生成 (RAG)

[ウ 検索拡張生成 (RAG) 等について]

素案では、AI 生成モデルが著作物を含む対象データを検索し、その結果の要約を生成する「検索拡張生成 (RAG)」と呼ばれる手法が説明されています。⁸素案では、特定の状況において、RAG の生成物に、学習のために「検索」されたデータの創作的表現の全部または一部が含まれる場合、そのデータの最初の収集は、著作物の「享受目的」であったとみなされる可能性があるとして示唆しています。素案においては、このような場合、AI 学習のためにそのような著作物を収集し利用することには、法 30 条の 4 は適用されない可能性があるとしています。

この点についても、素案において学習データと AI システムの生成物を混同しないよう明確に記載することを求めます。AI システムが著作権を侵害する生成物を生成するのであれば、そのような侵害に対する救済措置はその生成物に適用されるべきです。当該著作物の学習のための利用に影響を及ぼすべきではありません。その利用は表現内容を「享受」するためではなく、著作権の保護対象ではない、当該データに関する統計的データを抽出するためであるからです。したがって、RAG に関する著作権上の懸念に関する評価は、生成物が著作物の著作権を侵害するかどうかに関心を当てるべきであり、学習のために利用される著作物や当該モデルの参照のために適用される技術的処理は、データ解析を目的としており、著作権侵害とはなりません。法第 47-5 の追加的例外は、条件が満たされる場合、RAG のために生成された生成物にも適用されるべきです。生成された生成物によって著作権が侵害されると判断された場合、情報解析目的のための複製に影響を及ぼすべきではありません。これは、学習データ、また、参照用に表形式に変換されたデータの両方に関してあてはまります。

商用データベース

[(ウ) 情報解析に活用できる形で整理したデータベースの著作物の例について]

素案では、権利者が商用データベースを通して情報解析に活用できるコンテンツのライセンスを提供する場合、法第 30 条の 4 の権利制限が商用データベースを通して提供されたコンテンツに適用されない可能性があることを示唆しています。⁹著作権法においては、著作権者がライセンスを提供するだけで、著作権の例外規定を覆すことが可能とならないことを、貴庁において明示することを我々は求めます。AI 学習に利用される個々の著作物の利用は、たとえデータベースが著作物の集合で構成されていたとしても、商用データベースの潜在的市場を必ずしも損なうものではありません。著作物が商用データベースに含まれているか否かにかかわらず、その目的が創作的表現の享受でない場合には、著作物の情報解析には著作権保護は及びません。

さらに、素案では、ファイル “robots.txt” への記述により、全ての AI 学習用クローラをブロックする措置が講じられていない場合であっても、ある AI 学習用クローラをブロックする措置が講じられている場合には、当該ウェブサイト内のデータを情報解析に活用できる形で整理したデータ

⁸ 素案 5. (1) ウ 検索拡張生成 (RAG) 等について

⁹ 素案 5. (1) エ (ウ) 情報解析に活用できる形で整理したデータベースの著作物の例について

ベースの著作物が将来販売される予定があることを推認させる一要素となると考えられると説明しています（23 頁、脚注 24）¹⁰。この推認の根拠となる経験則や証拠を採用した裁判例や調査結果が素案には示されていないことを踏まえると、事実上の推認の根拠となるような指針を貴庁が示すことは適当ではありません。データベースの著作物が将来販売される予定があることは、当該データベースの著作物が具体的に主張立証すべきです。

また、クローラのブロックは、単にセキュリティの確保やデータやコンテンツの保護等の観点から行われていることも多く、将来の販路が阻害される可能性を評価することは困難であり、また、AI の生成物が、学習用データセットに含まれる著作物と共通の創作的表現を有するか否かを予見することも困難です。このような観点から、法第 30 条の 4 のただし書の適用について、上記のような広範な解釈を行うことは、今後の AI 開発を阻害することにもなりかねず、避けるべきです。

学習済みモデルの廃棄請求

【侵害に対する措置について /カ（イ）学習済みモデルの廃棄請求について】

素案においては、ある状況下において、学習済みモデルが「侵害の行為を組成した物」、「侵害の行為によって作成された物」、又は「専ら侵害の行為に供された機械若しくは器具」であるとみなされる場合、その廃棄請求が認められる可能性があることが記されています。素案で示唆されているように、¹¹ AI モデルの開発には相当な時間と費用が費やされます。さらに、AI モデルは「侵害の行為を組成した物」でも、侵害行為によって「作成された物」でもなく、侵害行為に「専ら供された機械若しくは器具」でもありません。AI モデルの廃棄が適切な救済策であると示すことは妥当ではありません。差止請求の焦点は、侵害しない実際のモデルではなく、AI システムの利用を享受することや、侵害生成物をさらに作成する行為、また、当該生成物を侵害的な方法で利用行することにあてるべきです。AI モデルは著作物の複製物の集合体ではないのです。

また、日本法では、一般的に、差止めは、金銭的な賠償では事後的に回復不能な状況に限って許容されると解されているところ、素案の脚注 29 及び 31 の記載¹²は、「学習データとして利用される著作物と類似性のある生成物を高確率で生成する状態」や、RAG 等で「享受目的と非享受目的が併存する場合」において、直ちに廃棄命令が下される可能性があるを読み取れるものであり、誤解を招きかねません。したがって、廃棄命令のような救済手段は、原告が回復不能な損害を被ることとなり、かつ、権利者における訴訟の焦点である著作物をモデルが侵害し続けるリスクを軽減する技術的手段が他にない場合のような、極端かつ思いも寄らない事態においてのみ利用し得ると明示すべきです。

生成・利用段階

依拠性の判断

【5. (2)イ（イ）② AI 利用者が既存の著作物を認識していなかったが、AI 学習用データに当該著作物が含まれる場合】

¹⁰ 素案 5. (1) エ（エ）脚注 24

¹¹ 素案 5. (1) カ（イ）学習済みモデルの廃棄請求について 脚注 31

¹² 素案 5. (1) カ（イ）学習済みモデルの廃棄請求について 脚注 29 および 31

素案では、「AI 利用者が既存の著作物（その表現内容）を認識していなかったが、当該生成 AI の開発・学習段階で当該著作物を学習していた場合については、客観的に当該著作物へのアクセスがあったと認められることから、当該生成 AI を利用し、当該著作物に類似した生成物が生成された場合は、通常、依拠性があったと推認され、著作権侵害になりうると考えられる」とされています（30 頁）¹³。AI システムは、著作物そのものではなく、著作物ではない情報からの推論に基づいて生成物を作成する可能性があるため、このような一般的な推認を貴庁において行うべきではないと考えます。さらに、素案 5.（2）イ（イ）で示された事例は、AI が生成したコンテンツが著作権侵害と見なされるかどうかの判断には役立ちますが、通常、AI 学習に利用される著作物はパラメータ化されるため、特定の生成された生成物が、学習データ内の数十億の著作物のうち、ある特定の著作物に依拠しているかどうかを判断することは極めて難しく、不可能とも言えます。したがって、生成された生成物が著作権を侵害しているかどうかを評価する際に、学習データへの「依拠」を当然視することは避けるべきです。

侵害行為の責任主体

【キ 侵害行為の責任主体について】

素案においては、AI の生成物が他者の著作物を侵害したとされる場合の責任の分担が論じられています（32-33 頁）。¹⁴また、AI 開発者や AI 導入者が著作権侵害の予防としていくつかの措置を講じるよう提案しており、その中には、既存の著作物と類似する生成物の生成を回避するための技術的措置¹⁵が含まれています。現行の著作権法に基づかない責任規定を貴庁において推奨することは避けるべきと考えます。また、このような考え方は、特にオープンソースで、基盤となる AI ツールやモデルをリリースする開発者の意欲を削ぎ、AI 研究を非常に高価なものにしてしまいます。我々は、権利者、アーティスト、AI 導入者が、侵害を制限しうる自主的なベストプラクティスについて議論することを奨励します。

その他の論点

学習データの開示

【コ 学習に用いた著作物等の開示が求められる場合について】

素案においては、AI 生成物の既存の著作物への依拠性を判断するために、学習データの開示が求められる場合が示されています（34 頁）¹⁶。開示を求めることは、AI システムの設計や利用に関する企業秘密や機密情報や専有情報の開示につながる可能性があり、AI テクノロジーへの投資を抑制することにつながります。AI 企業にデータセットのコンテンツを開示させることは、事実上、貴重、もしくは極秘の企業秘密の公開を強制することになります。AI 開発者が学習データのソースの概要を開示することが適切な状況もあるかもしれませんが、詳細な説明は非現実的であり、要求すべきではないと考えます。日本の AI とクリエイティブ産業の競争力を高めるために

¹³ 素案 5.（2）イ（イ）② Ai 利用者が既存の著作物を認識していなかったが、AI 学習用データに当該著作物が含まれる場合

¹⁴ 素案 5.（2）キ 侵害行為の責任主体について

¹⁵ 素案 5.（2）キ ④侵害行為の責任主体について

¹⁶ 素案 5.（2）コ 学習に用いた著作物等の開示が求められる場合について

も、正当な権利者とイノベーターに技術中立的な保護を提供する著作権法の中核的な規定を維持することを推奨します。

生成物の著作物性

【(3) 生成物の著作物性について】

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利用された機器や技術に関わらず、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 は創造性を強化します。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。ソフトウェア開発では、プログラマーの指示に基づいてソフトウェアコードを生成する補助に使われています。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。AI が生成した作品が創作的要素を含まず、人間が作成した作品と組み合わせられている場合、組み合わせられた作品全体が保護されないとすべきではありません。保護されない AI の生成部分の著作物性は否認されたとしても、保護可能な作品の部分は著作権で保護されるべきです。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、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、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。いずれの場合も、人間の利用者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 が創作的表現の一部ではない場合もあるでしょう。

その他の論点

著作権者等への対価還元

【(4) その他の論点について】

素案では、著作物が AI の学習に利用された著作権者等への対価還元の手段についての考え方が示されています。¹⁷ AI の学習・開発のための情報解析の用に供するために著作物を利用することに

¹⁷ 素案 5. (4)

より、著作権者等の利益が通常害されるものではないため、対価還元的手段として、著作権法において補償金制度を導入することは理論的な説明が困難であるとする見解に我々は同意します。AI時代の知的財産権検討会（第5回）の「残された論点等（討議用）」¹⁸でも、「一律の収益還元策を講じることは適切とはいえ、本検討会としては、民間において講じることが考えられる任意の方策例について、提示することが適切ではないか」との考え方が示されており、貴庁の素案においても同様のアプローチを採るべきです。

結論

BSAは、貴庁に意見を述べた機会を得られたことに感謝します。侵害のリスクを最小化する取り組みにおいて、AIの学習プロセスに関する理解を促進するために、多様な関係者を参画させることを奨励します。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。さらに、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に検討する取り組みをBSAは支持します。今後は、貴庁において、意見募集に十分な期間を設けることを希望します。今回、BSAや他の関係者に、この重要な文書を翻訳、検討、協議するための、適切な期間が与えられなかったことを残念に思います。1月23日に公表され、2月12日が意見提出期限となっていますが、これでは関係者が対応するのに3週間以下の期間しか残されておらず、このような短期間で、定められた時間内に思慮深く建設的なフィードバックを提供することは非常に困難です。関係者が実質的で、思慮深く、建設的な意見を提供できるよう、このような詳細な文書については、少なくとも4週間、理想的にはそれ以上の期間を設けて頂ければ幸いです。

¹⁸ AI時代の知的財産権検討会（第5回）の「残された論点等（討議用）」
https://www.kantei.go.jp/jp/singi/titeki2/ai_kentoukai/gijisidai/dai5/siryou1.pdf