

AI・機械学習と個人の権利 —著作権法と個人情報保護法の横断的考察

栗田 昌裕¹ (名古屋大学)

要 旨

AIの開発・学習段階では、①学習用データの収集、②学習用データセットへの加工及び③機械学習による学習済みモデルの構築の各段階において、情報の利用が行われる。このとき、利用される情報に著作物又は要配慮個人情報等が含まれていれば、著作権者の許諾又は本人の事前同意を取得しなければならない。しかし、インターネットを介して公にされている情報をクロールやスクレイピングによって収集又は抽出する場合において、各情報の著作物性や要配慮個人情報該当性を判断し、著作権者や本人を特定したうえで、その許諾又は事前同意を取得しなければならないとすれば、AIの開発及び学習に禁止的な費用を課すことになりかねない。また、権利者等の意思によって特定の情報を学習の対象から除外できるとすれば、結果として偽情報の拡散を助長するおそれがあるほか、AIシステムの精度が低下し、ハルシネーションの生じる確率が上がることにより、社会全体の便益を損ないかねない。さらに、個人情報保護法との関係では、インターネットを介した情報の利用について頻回に詳細な同意を求めることが、かえって同意の形骸化を招き、個人の自律を支援する意義を失わせるおそれがあるとも指摘されている。もっとも、著作権者が著作物の利用行為を禁止することができ、交渉して対価を取り決めることができるのは、創作者の経済的自律を保障するための重要な法的手段であり、財産権保障の対象でもある。また、本人が自らの個人情報等の取扱いについて同意を与える権利を留保することは、その人格的自律にとって重要な意味を有する。そうだとすれば、情報の自由な利用と個人の権利との間に適切な調整が図られるように、法解釈と立法を行うことが望まれる。

本稿では、こうした観点から、開発・学習段階における情報の利用について、著作権法と個人情報保護法の分野横断的な考察を試みた。その際には、主として著作権法30条の4第2号及び個人情報保護法20条各項の解釈を取り上げた。両法を考察の対象とするのは、いずれかの法領域において著作権者の許諾又は本人の事前同意の取得が要求されれば、結果として情報の自由な利用が抑制されるからである。一方の解釈が直ちに他方に影響するものではないが、利害状況の比較衡量にあたっては、共通する背景を有することに意を払う必要がある。また、情報の自由な利用と個人の権利の調整という観点からは、情報の流通によって侵害される権利としての民法上のプライバシーにも言及しておく必要があるだろう。プライバシーの保護は表現の自由等に対して抑制的に作用し得るからである。そこで、本稿では、判例法を中心として、民法上のプライバシーにも簡単に触れた。

キーワード：AI、機械学習、情報解析、柔軟な権利制限規定、個人情報、プライバシー

¹ 名古屋大学大学院法学研究科教授。

1. はじめに

1. 1. 本稿の目的

本稿では、AI システムの開発・学習段階における情報の自由な利用と個人の権利の調整が問題となる法制度について、分野横断的な考察を行う。具体的には、著作権法を中心としつつ、個人情報保護法（個人情報の保護に関する法律）及び民法上のプライバシーにも必要な範囲で言及する。

1. 2. 機械学習と情報の利用

AI システムの開発・学習段階では、①学習用データ（生データ）の収集、②学習用データから学習用データセットへの加工及び③学習用データセットの学習前モデルへの入力（機械学習）の各段階において、情報の利用が行われる。このうち、①学習用データの収集には、書籍、雑誌、新聞等の媒体からの入力のほか、ウェブクロールやウェブスクレイピング等（以下「クロール等」という。）のウェブサイトを自動的に巡回して情報を収集又は抽出する技術が用いられることがある。次に、②学習用データ（生データ）から学習用データセットを構築するために、ノイズ除去等（クレンジング）、数値の標準化（スケーリング）、機械学習可能な形式への変換（エンコーディング）及びラベルの付加（アノテーション）等の前処理が行われる。最後に、③学習用データセットの学習前モデルへの入力（機械学習）によって学習済みモデル（学習済みパラメータ）が生成される。また、追加学習では、その目的によって学習されるデータの範囲は異なるものの、基本的には学習済みモデルについて①から③までの各過程が実行されることになる。そのほか、検索拡張生成（Retrieval-Augmented Generation: RAG）では、生成される応答結果の精度等を向上させるために、学習済みモデルに外部データベースが接続される。この場合には、外部データベースの構築と応答結果の出力に際して情報の利用が行われる²。

1. 3. 情報の利用と法規制

開発・学習段階において利用されるデータには、著作物や個人情報等が含まれることがある。その場合には、著作物の利用行為として著作権法の適用を受け、又は個人情報の取扱いとして個人情報保護法の規制を受ける。しかし、クロール等による情報収集に際して、常に著作権者の許諾を受け、又は本人の事前同意を取得しなければならないとすれば³、禁止的な費用の増加をまねきかねない。また、学習済みモデルに偏向（バイアス）を生じさせるおそれもある。たとえば、対話型生成 AI システム等に用いられる大規模言語モデル（Large Language Model: LLM）の開発には、大量の言語資料から構築されたコーパスが必要とされる。このとき、収集又は抽出された言語資料のすべてについて著作物性や要配慮個人情報該当性を判断し、著作権者や本人を特定して、その許諾や事前同意を得なければならないとすれば、コーパスの構築に要する費用は増大し、又は許諾や事前同意を得られたも

² 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）18頁以下参照。

³ 現行法上、その取得にあたって本人の事前同意が必要とされるのは要配慮個人情報に限られる（個人情報保護法20条2項柱書）。

のに学習の対象が偏ることによって、学習済みモデルの精度が失われるおそれがある⁴。しかし、コーパスの構築は、著作物に表現されている思想又は感情の享受を目的としているわけではないし、取得した（要配慮）個人情報の利用を目的としているわけでもない。そのため、学習用データの収集又は抽出そのものについて、著作権の侵害又は個人情報保護法違反として直ちに違法評価をすべきかが問題となる⁵。また、期待可能性のない義務を賦課すれば、かえって違法な情報の利用を助長しかねない。

もっとも、著作権者が著作物の利用行為を禁止することができ、交渉して対価を取り決めることができるのは、創作者の経済的自律を保障するための重要な法的手段であり、憲法上の財産権保障の対象でもある（憲法 29 条 1 項）。また、本人が自らの個人情報等の取扱いについて同意を与える権利を留保することは——これを自己情報コントロール権と呼ぶかどうかにかかわらず——その人格的自律にとって重要な意味を有する（憲法 13 条参照）。そのため、これらの自律に対する制約は最小限にとどめなければならず、可能な範囲で自律を支援することが求められる。

しかし、AI システムの開発・学習段階から生成・利用段階まで、情報の利用は数次にわたって行われる。このうち、クロージング等による情報収集に際して、偶発的に生じる著作物の利用や個人情報等の取得を個人の意思に係らしめることが、その自律にとって真に重要な意義を有するのか改めて問い直す必要がある。たしかに、既存の著作物に類似した生成物が大量に出力されることによりクリエイターの利益が害されたり、AI システムによる個人情報等の利用によって、プロファイリングによる機微情報の推知、ターゲティング広告による脆弱性の利用、パーソナライズドプライシングによる消費者余剰の搾取又はスコアリングを介した差別や偏見の固定化等の弊害が生じたりすることが指摘されている。しかし、これらの弊害は、主として出力・利用段階において生じるものである。その抑止は、必ずしも開発・学習段階における情報の利活用を個人の意思に係らしめることによらなければ実現し得ないわけではない。

1. 4. 本稿の視角

このように、開発・学習段階における情報の利用を巡っては、情報の自由な利用の要請と個人の自律とが緊張関係にある⁶。もっとも、すべての情報の利用について個別に同意の取得を義務づけるとすれば、事業者には禁止的な費用の増加や学習済みモデルの偏向を、個人

⁴ 前川喜久雄「コーパス構築と著作権保護」人工知能学会誌 25 巻 5 号(2010)1 頁参照。ただし、学術研究機関等が学術研究目的で要配慮個人情報を取得する場合には、本人の事前同意は必要とされない（個人情報保護法 20 条 2 項 5 号）。

⁵ 著作権の侵害については、侵害組成物廃棄請求権を行使することができ（同法 112 条 1 項）、罰則規定も設けられているため（同法 119 条 1 項）、より営業の自由にとって制約的である。

⁶ AI システムの社会実装によって社会全体の厚生が増進が期待されることに鑑みれば、この問題を功利主義と権利論の対立と理解することができるかもしれないが、本稿では基底的価値の対立には立ち入らない。なお、著作権者は個人とは限らないが、その原始的帰属主体である著作者は職務著作や映画の著作物の例外を除いて個人であるため（著作権法 15 条、同法 16 条、同法 29 条 1 項参照）、ここでは原則的事例に着目し、個人の自律の問題として考える。

にはいわゆる「同意疲れ」「同意の形骸化」⁷を、それぞれもたらしかねない。そのため、事業者に課すべき義務の期待可能性と、個人の同意の実質化を考慮した制度設計が求められる。

以下では、こうした視点から、開発・学習段階における情報の利用について、著作権法と個人情報保護法の分野横断的な考察を試みる。両者を考察の対象とするのは、いずれかの法領域において著作権者の許諾又は本人の事前同意の取得が要求されれば、いずれにせよクローリング等による大規模かつ網羅的な情報の収集は妨げられるからである。なお、両法領域の規律が及ばない分野について不法行為法等が法的救済を与えることがあるため、必要に応じて民法の規律にも言及する。

2. 著作権法と機械学習

2. 1. 序

既存の著作物を学習用データ（生データ）として収集する行為と、これを学習用データセットへと加工する行為は、いずれも複製に当たる。また、機械学習の方式にもよるが⁸、学習前モデルに学習用データセットを入力する際にも複製が行われることが多い。著作権法21条は、「著作者は、その著作物を複製する権利を専有する」と規定しており、原則として、これらの利用行為には著作権者の許諾を得なければならない。ただし、著作権の制限規定が適用されれば、無許諾で自由に利用することができる。AI システムの開発・学習段階における著作物の利用行為については、それが著作権法30条の4第2号による権利制限の適用対象となるかどうかの主として論じられている。

2. 2. 著作権法30条の4第2号（非享受目的利用）

著作権法30条の4柱書本文は、「著作物は、次に掲げる場合その他の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場合には、その必要と認められる限度において、いずれの方法によるかを問わず、利用することができる」と規定し、同条2号は「情報解析（……）の用に供する場合」を例示している。

同条は、平成30年著作権法改正（平成30年法律第30号）における「デジタル化・ネットワーク化の進展に対応した柔軟な権利制限規定の整備」の一環であり、同年改正前47条の7（情報解析のための複製等）を基礎としつつ、①「電子計算機による情報解析」（強調筆者。以下同じ）という要件から「電子計算機による」との限定を、②「多数の著作物その他の大量の情報から、当該情報を構成する言語、音、影像その他の要素に係る情報を抽出し、比較、分類その他の統計的な解析を行うこと」という情報解析の定義から「統計的な」という限定をそれぞれ削除し、③権利制限の対象を「記録媒体への記録又は翻案」から著作物の利用一般へと拡大したものである⁹。こうした経緯から分かるように、同法30条の4第2

⁷ 総務省「プラットフォームサービスに関する研究会最終報告書」（2020年2月）12頁以下、同「同意取得の在り方に関する参照文書」14頁以下参照。

⁸ 前掲「AIと著作権に関する考え方について」18頁。たとえば、ストリームデータ処理を用いたインクリメンタル学習やクライアントからサーバに学習済みモデルのみを送信して統合する連合学習（Federation Learning）等がある。

⁹ 文化庁著作権課「デジタル化・ネットワーク化の進展に対応した柔軟な権利制限規定に

号は、当初から、「深層学習（ディープラーニング）の方法による人工知能の開発のための学習用データとして著作物をデータベースに記録するような場合」への適用を想定していた¹⁰。

2. 2. 1. 情報解析と非享受目的要件

もっとも、著作権法 30 条の 4 柱書本文の「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場合」という要件（以下「非享受目的要件」という。）と同条各号との関係については争いがある。

まず、同条 2 号の要件を満たせば、常に「非享受目的」と評価すべきとの見解が主張されている。この見解は、同条柱書本文の「次に掲げる場合その他の」という文言から同条各号は非享受目的利用の例示と解釈すべきことと、非享受目的を固有の要件とすれば同条各号の適用を主張立証する意味が没却されることをその理由とする¹¹。

これに対して、同条柱書本文の適用は「自ら享受し又は他人に享受させることを目的としない場合」に限定されており、非享受目的と享受目的が併存する場合には、同条各号に該当する場合であっても同条を適用すべきではないとの指摘がある¹²。この見解を採れば、同条 2 号（情報解析の用に供する場合）に該当するだけでなく、同条柱書本文の非享受目的要件をも満たさなければ、同条の権利制限は適用されないことになる¹³。同条 1 号の解釈との均衡に鑑みて、後者の見解を支持したい。

2. 2. 2. 享受目的の併存

この見解を採るとき、享受目的の存否を判断するにあたって生成・利用段階における目的も考慮すべきかどうかについては、立場が分かれている¹⁴。これを肯定する立場からは、生成・利用段階における著作権の侵害を未然に防止すべきこと¹⁵、ただし、開発・学習段階において学習の対象とされた著作物（以下「学習対象著作物」という。）の表現が出力されな

関する基本的な考え方について（著作権法第 30 条の 4、第 47 条の 4 及び第 47 条の 5 関係）（令和元年 10 月 24 日）10 頁（以下「基本的な考え方」という。）、同「著作権法の一部を改正する法律（平成 30 年改正）について（解説）」24 頁（以下「平成 30 年改正解説」という。）、加戸守行『著作権法逐条講義〔七訂新版〕』（著作権情報センター・2021）286 頁以下。このような規定の趣旨からすれば、学習用データセットの構築に当たって人の手でラベル付加を行う場合にも、本号は適用され得ることになる。

¹⁰ 基本的な考え方 10 頁、平成 30 年改正解説 24 頁、加戸・前掲注 9）286 頁。

¹¹ 愛知靖之「日本法における権利制限」上野達弘＝奥邨弘司監修『AI と著作権』（勁草書房・2024）16 頁以下、上野達弘ほか「座談会」同書 208 頁以下、218 頁以下〔愛知発言〕（以下「座談会」という）。

¹² 座談会 209 頁以下〔奥邨発言〕、同 211 頁以下〔前田健発言〕、同 215 頁〔横山久芳発言〕、同 216 頁〔今村哲也発言〕。

¹³ 著作権法 30 条の 4 第 2 号の「情報解析」の限定解釈との整理を示唆するものとして、座談会 211 頁〔上野発言〕。

¹⁴ 肯定説として、座談会 212 頁〔前田発言〕、同 213 頁以下〔奥邨発言〕、同 215 頁以下〔横山発言〕、同 216 頁〔今村発言〕。否定説として、座談会 213 頁〔谷川和幸発言〕。

¹⁵ 座談会 215 頁以下〔横山発言〕。

いよう調整がされていた場合には享受目的を否定すべきこと¹⁶、及び既存の著作物の創作的表現に類似する生成物（以下「侵害生成物」という。）の出力の頻度等も含めて享受目的の有無を判断すべきことなどが指摘されている¹⁷。

2. 2. 3. 通貫目的基準

また、非享受目的要件との関係では、「学習対象著作物の表現が、AI 内部に保存されて、（少なくともその一部が）出力されること（「通貫」と呼ぶ）を目的とするか否か」を著作権法 30 条の 4 の適用があるかどうかの判断基準とする見解が主張されている。この立場では、通貫目的のある場合には、学習対象著作物の享受目的利用が併存するため、同条は適用されず、同法 47 条の 5 の問題となる。通貫目的のない場合には、同法 30 条の 4 の問題となる。このように、通貫目的の有無によって、同法 30 条の 4 と同法 47 条の 5 の適用範囲が明確に区別されることになる。なお、この立場は、「作風や画風のような、著作権で保護されないもの」が共通するとしても、同法 30 条の 4 柱書ただし書を適用すべきではないとする。このような行為を著作権者の利益を「不当」に害するものと評価すべきではなく、推論過程の問題を学習の禁止によって解決しようとするのは得策ではないからである¹⁸。

2. 2. 4. 分析と検討

情報解析の用に供するための著作物の利用行為は、原則として非享受目的であり、著作権法 30 条の 4 第 2 号の適用を受ける。ただし、例外的に享受目的が併存する場合には、同条柱書本文の要件を満たさないため、権利制限の対象とならない。

そのうえで、本稿では、著作物の利用行為の時点で、その著作物の創作的表現に類似する生成物の出力を目的としていれば、例外的に享受目的が併存するものとする。その目的を達成するための手段や過程として何を想定していたかは問わず、生成・利用過程において現に侵害生成物が出力されたことも要しない。ただし、享受目的の併存を認めることができるのは、特定の著作物に類似する生成物の出力を「目的」としてその著作物の利用行為（学習過程における複製等）がされた場合に限られる¹⁹。

このとき、享受目的の存在は、著作物の利用行為時を基準として、その利用行為について判断されなければならない。したがって、行為時、作風や画風が共通するにとどまる——学習対象著作物との類似までは認められない——生成物の出力を目的としていた場合には、享受目的の併存は認められない。また、学習対象著作物に類似する生成物が出力され得ると

¹⁶ 前掲「AI と著作権に関する考え方について」20 頁脚注 22、座談会 218 頁〔前田発言〕。

¹⁷ 座談会 220 頁以下〔奥邨発言〕。同発言は、開発・学習段階において学習対象著作物の表現が出力されないよう試行錯誤する際の著作物の利用行為には著作権法 30 条の 4 第 1 号が適用され、実装後が同条 2 号の問題になると整理する。

¹⁸ 奥邨弘司「生成 AI と著作権に関する米国の動き」コピーライト 747 号(2023) 46 頁以下、50 頁注 60、50 頁以下注 64。

¹⁹ たとえば、「ジブリ風の画像が生成できるモデルを作った」というだけでは、享受目的が存在するとはいえない（座談会 223 頁以下〔谷川発言〕参照）。これに対して、開発・学習段階において、学習対象著作物である特定のカットに類似する画像を生成する目的があれば、享受目的の併存を認めてよい。

いう抽象的かつ一般的な可能性を単に認識していたに過ぎない場合も、享受目的の併存は認められない。「目的」は、故意過失とは区別されるべき主観的要件だからである²⁰。そのほか、享受目的の不存在を強く推認させる事情として、著作物の利用者（AI 開発者等²¹）が侵害生成物の出力を抑止するための技術的手段を講じていることが挙げられる²²。これは、学習対象著作物の表現が享受されることを阻止するという意図の徴表だからである。

これに対して、生成・利用段階において侵害生成物が出力される頻度等の事情から直ちに開発・学習段階における享受目的の併存を認定することには慎重であるべきだろう²³。AI 提供者及び AI 利用者の行為が介在する生成・利用段階の帰結は、必ずしも開発・学習段階における AI 開発者の「目的」と一致するとは限らないからである。もっとも、有名作品のキャラクター名をプロンプトとして入力するだけで高頻度で学習対象著作物に類似する生成物が出力されるような場合であって、その分野における通常の知識を有する者であれば技術的措置等によって容易にこれを防止できたときには、これらの事情を開発・学習段階における享受目的の併存を推認させるものとして考慮してよいだろう。

また、追加学習等にあって、侵害生成物を出力させる目的で著作物の複製を行う場合には、享受目的が併存するものと評価できる²⁴。たとえば、学習対象著作物に類似する生成物を出力させることを目的として、低ランク適応（Low-Rank Adaptation: LoRA）を行う場合がこれに当たる²⁵。この場合には、著作物の利用行為の時点において「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」としてその利用行為をしたといえるからである。

このように、特定の著作物に類似する生成物を出力させることを「目的」とするのであれば、開発・学習段階において学習対象著作物の著作権者から許諾を得るように求めても、AI 開発者等に期待可能性のない義務を課すことにはならない。著作権者の許諾が得られないために、その著作物に類似する生成物を出力させる AI システムの開発が行えないとしても、これは、原著作物の著作権者の許諾が得られないために、その二次的著作物を出版できない

²⁰ ただし、例外的とはいえない範囲の者が著作権侵害に利用する蓋然性が高いと認められ、これを認識、認容しながら生成 AI サービスの公開、提供を行った場合には、幫助に該当し得ると指摘されており（座談会 302 頁〔谷川発言、奥邨発言〕。最決平成 23 年 12 月 19 日刑集 65 卷 9 号 1380 頁参照）、これは別論である。

²¹ AI 開発者（AI Developer）、AI 提供者（AI Provider）及び AI 利用者（AI Business Use）の区別については、総務省＝経済参照省「AI 事業者ガイドライン（第 1.1 版）」（令和 7 年 3 月 28 日）5 頁以下参照。

²² AI 提供者が技術的措置（侵害防止措置）を何らとっていなければ、出力の主体とされ、違法評価を受けるという指摘として、座談会 299 頁〔横山発言〕。技術的措置をとったかどうかにかかわらず、侵害生成物が出力されれば、開発・学習段階における著作物の利用行為に違法性を認める見解として、座談会 300 頁〔愛知発言〕。

²³ 類似性の高い生成物が出力される頻度やそのための試行回数を依拠性の要件において考慮する見解として、奥邨弘司「依拠・類似」前掲『AI と著作権法』125 頁、座談会 296 頁〔奥邨発言〕。

²⁴ 前掲「AI と著作権に関する考え方について」20 頁、24 頁。

²⁵ 低ランク適応（LoRA）の著作権法上の評価について、座談会 245 頁以下〔奥邨発言〕。同発言は、同法 47 条の 5 の適用可能性を論じる文脈において行われたものだが、付随的に同法 30 条の 4 の適用可能性にも言及する。

のと質的に変わらないからである。また、クローリング等によるコーパス構築の例とは異なり、ここでは許諾を得るべき著作権者の数や特定の困難性を理由として著作権の制限を基礎づけることもできない。これらの事情は、学習対象著作物の選択にあたって事業者が考慮すべき営業判断の一要素に過ぎないからである²⁶。さらに、ここでは、侵害生成物を通じて「[学習対象] 著作物に表現された思想又は感情」を享受できることから経済的利益を得ようとしているのであって、その開発・学習段階における複製等は、著作権によって本来的に捕捉されるべき利用行為だといえる。

以上のように、非享受目的要件の解釈にあたって、開発・学習段階における情報の自由な利用と著作権者の利益との調整が考慮される。

2. 3. 著作権法 30 条の 4 柱書ただし書

著作権法 30 条の 4 柱書ただし書では、「当該著作物の種類及び用途並びに当該利用の態様に照らし著作権者の利益を不当に害することとなる場合」には、同条の権利制限の適用を受けないことを定めている。同条ただし書に該当するかどうかは、「著作権者の著作物の利用市場と衝突するか、あるいは将来における著作物の潜在的販路を阻害するかという観点」から、個別具体的に判断される²⁷。

これに対して、著作権法 30 条の 4 第 2 号の前身である平成 30 年改正前著作権法 47 条の 7 は、そのただし書において、「情報解析を行う者の用に供するために作成されたデータベースの著作物」のみを適用対象から除外していた。そのため、現行法でも、著作権法 30 条の 4 柱書ただし書が「大量の情報を容易に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が販売されている場合に、当該データベースを情報解析目的で複製等する行為」に適用されることには争いがない²⁸。たとえば、オンラインで学習用データセットやデータの取得を可能にする API (Application Programming Interface) が有償で販売等され、または販売等が予定されている場合は、これに当たる²⁹。

もっとも、現行著作権法 30 条の 4 柱書ただし書には——改正前著作権法 47 条の 7 とは異なり——その適用をデータベースの著作物に限定する文言はない。そのため、AI との関係では、『作風』や『画風』といったアイデア等が類似するにとどまり、既存の著作物との類似性が認められない生成物が大量に出力されるなど、「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場合」にも同条た

²⁶ 低ランク適応 (LoRA) 等を想定すれば、著作権者は最少 1 名、学習対象著作物は数十件程度であるため、いずれも相対的に少数と予想される。AI 開発者は、著作権者を特定しその許諾を得られる限りにおいて、多数の著作権者に係る多数の著作物を学習の対象としてもよく、許諾等に要する費用は営業判断の一要素といえる。ここで重要なのは、必要とされる学習対象著作物の件数が少数であるため、著作物性の判断や著作権者の特定等に要する費用が常に過大になるとは限らない、ということである。

²⁷ 基本的な考え方 9 頁、平成 30 年改正解説 25 頁、加戸・前掲注 9) 284 頁、前掲「AI と著作権に関する考え方について」23 頁。

²⁸ 平成 30 年改正解説 25 頁、加戸・前掲注 9) 284 頁、前掲「AI と著作権に関する考え方について」24 頁。

²⁹ 前掲「AI と著作権に関する考え方について」24 頁以下。

だし書を適用すべきかどうかについて争いがある³⁰。

2. 3. 1. 「作風」「画風」の共通する AI 生成物による市場競合を認める見解

第一に、「特定著作権者の著作物を狙い撃ちし、当該著作権者の『作風』を備えることで市場競合をもたらすコンテンツを生成するために、その著作物の全てあるいはこれを大量に学習・推論に利用する行為は、〔著作権法 30 条の 4 柱書〕但書に該当する」という立場がある³¹。この立場は、同条の立法趣旨は「原則として、非享受利用は著作物の享受利用市場と衝突することはないこと」を前提としているところ、前記の行為は「潜在的な著作物の享受利用市場」との衝突を引き起こすため、権利制限の対象とする必要はないと指摘する。また、作風や画風はアイデアであって著作物ではなく、これが共通する AI 生成物の出力は著作物の利用には当たらないとの指摘に対しては³²、ここで問題となるのは「AI に入力＝複製するという利用行為」であるため、たとえ「AI 内部で入力された著作物からアイデアのみを抽出するというプロセスが取られたとしても、AI への著作物の複製という利用行為が行われていたことに変わりはない」とされる³³。

2. 3. 2. 通貫目的を基準とする見解

第二に、すでに見たように通貫目的の有無を基準として同法 30 条の 4 と同法 47 条の 5 の適用場面を区別する見解がある。この見解では、学習対象著作物の表現を出力することを目的とする場合には（通貫目的）、享受目的が併存するために同法 30 条の 4 は適用されない。これに対して、通貫目的が存在しない場合には、「作風や画風のような、著作権で保護されないもの」が共通するとしても、同法 30 条の 4 柱書ただし書を適用すべきではなく、権利制限の対象とされる。

2. 3. 3. 非享受目的市場の競合を要件とする見解

第三に、同法 30 条の 4 柱書ただし書の適用を「非享受利用されることを主たる目的として制作された著作物を非享受目的で利用する行為に限定される」とする立場や³⁴、これに加えて「非享受目的のライセンス市場が発展し当該著作物について非享受目的の利用が本来的な利用と客観的に評価できるに至った場合には、そのライセンス市場と衝突するような利用については、ただし書に該当することもあり得る」とする立場がある³⁵。ただし、後者の見解については、取引慣行としてライセンス市場が成立することは非侵害行為を侵害行

³⁰ 前掲「AI と著作権に関する考え方について」23 頁参照。

³¹ 愛知・前掲注 11)28 頁以下、同「AI 生成物・機械学習と著作権法」パテント 73 巻 8 号(2020)142 頁以下。齋藤浩貴「エンタテインメント分野における生成 AI の利用と著作権」コピライト 752 号(2023)12 頁以下、山崎貴啓「生成 AI と著作権」コピライト 754 号(2024)56 頁参照。

³² 奥邨・前掲注 18) 50 頁以下注 64 参照。

³³ 愛知・前掲注 11) 29 頁脚注 38。

³⁴ 柿沼太一「生成 AI と著作権」知財ぷりずむ 21 巻 248 号(2023)22 頁。

³⁵ 松田政行編『著作権法コンメンタール別冊 平成 30 年・令和 2 年改正解説』（勁草書房・2022）32 頁〔澤田将史〕。

為に転化させる理由にならないとの批判がある³⁶。

2. 3. 4. 平成30年改正前著作権法との連続性を重視する見解

第四に、平成30年改正に係る附帯決議において「柔軟な権利制限規定の導入に当たっては、現行法において権利制限の対象として想定されていた行為については引き続き権利制限の対象とする立法趣旨を積極的に広報・周知すること」と指摘されていることから³⁷、少なくとも旧法において権利制限の対象とされていた「記録媒体への記録又は翻案」については、データベースの著作物の複製に限って著作権法30条の4柱書ただし書の適用を認めるべきであるという解釈の可能性が指摘されている³⁸。もっとも、前述のとおり、現行法の文言にはその適用の対象を特定の種類の著作物に限定する文言はなく、文理上、前記のような限定解釈が必然的に要請されるわけではない³⁹。その解釈にあたっては、より実質的な検討が必要であろう。

2. 3. 5. 分析と検討

以上のように、著作権法30条の4柱書ただし書の解釈を巡っては、学説の対立がある。このうち、「特定著作権者の著作物を狙い撃ちし、当該著作権者の『作風』を備えることで市場競争をもたらすコンテンツを生成するために、その著作物の全てあるいはこれを大量に学習・推論に利用する行為」への適用を認める見解には、以下の疑問がある。

たしかに「作風」や「画風」（以下「作風等」という。）の共通するAI生成物が出力されることにより、既存の著作物の潜在的販路が阻害されるおそれはある。しかし、著作権法による保護の対象は未創作の著作物に及ぶものではなく、クリエイターに対する需要の代替を理由として著作権の侵害を認めるべきではない。

次に、作風等はアイデアに属するところ、著作権法はアイデアを保護するものではなく、作風等を独占することを「著作権者の利益」ということはできない。そのため、作風等が類似するにとどまる生成物の出力を目的とすることから、情報解析の用に供するための著作物の利用を「著作権者の利益を不当に害する」と評価すべきではない⁴⁰。アイデアの共通する著作物によって既存の著作物の利用市場との衝突やその潜在的販路の阻害が生じることは、著作権法にすでに織込み済みだといえる。

この点については、作風等には「特定の作者固有のオリジナリティが認められ、その者の

³⁶ 愛知・前掲注11)27頁以下。柿沼・前掲注34)22頁参照。著作権の侵害が刑事罰の対象となり得ることも考えれば、この批判は正当と思われる。

³⁷ 参議院文教科学委員会「著作権法の一部を改正する法律案に対する附帯決議」（平成30年5月17日）。基本的な考え方9頁参照。

³⁸ 上野達弘「平成30年著作権法改正について」高林龍ほか編『年報知的財産法2018-2019』（日本評論社・2018）5頁注8、同「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(2019)39頁、同脚注34、同「情報解析と著作権——『機械学習パラダイス』としての日本」人工知能36巻6号(2021)747頁、座談会226頁〔上野発言〕参照。

³⁹ 愛知・前掲注31)143頁、同・前掲注11)25頁以下。

⁴⁰ 奥邨・前掲注18)47頁、座談会223頁以下〔谷川発言〕、224頁以下、235頁以下〔奥邨発言〕、225頁以下〔前田発言〕、235頁以下〔奥邨発言〕。

作品であると容易に同定し得るような具体性及び表現との近接性・連続性を備えた」「同一著作権者の（複数作品にまたがる）諸要素の一体的結合」と評価できるものがあり、この意味での作風等を間接的に保護することになっても、著作物の自由利用に対する過度の制約がもたらされるわけではないという反論がある⁴¹。たしかに、アイデアと表現は必ずしも明確に区別できるわけではない。しかし、そうであっても、作風等は経時的にも共時的にも複数のクリエイターが相互に影響を与え合って成立するものであるから、その独占を認めるべきではない。また、この見解は、特定の著作物の創作的表現ではなく、クリエイターの固有のオリジナリティを基準としているため、複数のクリエイターが類似する作風等を用いたり、1人のクリエイターが複数の作風等を使い分けていたりする場合には、被侵害者の特定が困難になるおそれがある。

そのほか、主張立証責任の分配の見地から、又は侵害生成物の出力による著作権者の不利益や営利目的と非営利目的の区別を考慮すべきことを理由として、同条ただし書の解釈として柔軟な判断を行うべきとも指摘されている⁴²。しかし、主張立証責任の分配の問題は、同法ただし書の解釈によらずとも解決することができる⁴³。行為規範としての明確性を維持する意味でも、同条ただし書の適用には慎重であるべきだろう。

以上のように、少なくとも現時点では、著作権法 30 条の 4 柱書ただし書の適用範囲を、解釈によって、作風等の共通する生成物の出力を目的とする AI システムの開発のための著作物の利用にまで拡張する必要はないと考える⁴⁴。

2. 4. 著作権法 47 条の 5 第 1 項 2 号（軽微利用等）

平成 30 年著作権法改正によって導入された「デジタル化・ネットワーク化の進展に対応した柔軟な権利制限規定」のうち、AI システムの開発・学習段階における著作物の利用への適用が考えられるものとして、著作権法 47 条の 5（電子計算機による情報処理及びその結果の提供に付随する軽微利用等）がある。もっとも、同条による権利制限には、利用主体に政令で定める基準の遵守が求められること、利用客体が公衆提供等著作物に限られること、付随性及び軽微性が要件とされることなどの限定がある。

まず、同条 1 項 1 号は、「電子計算機を用いて、検索情報が記録された著作物の題号又は著作者名、送信可能化された検索情報に係る送信元識別符号その他の検索情報の特定又は所在に関する情報を検索し、及びその結果を提供すること」（所在検索サービス）に付随して行われる軽微利用を権利制限の対象とするものである。その典型例としては、「特定のキーワードを含む書籍を検索し、その書誌情報や所在に関する情報の提供に付随して、書籍中の当該キーワードを含む文章の一部分を提供する行為（書籍検索サービス）」等が想定されている⁴⁵。

⁴¹ 愛知・前掲注 11) 32 頁以下。

⁴² 座談会 232 頁以下〔愛知発言〕、233 頁〔横山発言〕。

⁴³ 座談会 233 頁以下〔前田発言〕。

⁴⁴ 前掲「AI と著作権に関する考え方について」24 頁以下参照。

⁴⁵ 基本的な考え方 20 頁、加戸・前掲注 9) 418 頁。対話型生成 AI の出力する応答結果に検索結果の URL とともにその内容のスニペットが含まれている場合を想定すればよい。検索拡張生成（RAG）を想定した議論として、座談会 246 頁以下〔今村発言、上野発

次に、同条1項2号は「電子計算機による情報解析を行い、及びその結果を提供すること」（情報解析サービス）に付随して行われる軽微利用を権利制限の対象とする。その典型例としては、「検証したい論文との文章の一致について解析を行い、他の論文等からの剽窃の有無や剽窃率に関する情報の提供に付随して、剽窃箇所に対応するオリジナルの論文等の本文の一部分を表示する行為（論文剽窃検証サービス）」等が想定されている⁴⁶。

また、これらの行為の準備のための情報の収集、整理及び提供を政令で定める基準に従って行う者は、その準備のために必要と認められる限度において、複製、公衆送信及び頒布を行うことができる（同2項）。同条1項各号の行為をAIシステムの利用により行う場合において、その開発・学習段階において著作物の複製等を行うときには、同条2項の権利制限が適用され得る⁴⁷。例えば、検索拡張生成における外部データベースの構築は、このような準備行為に該当し得るだろう。

もっとも、生成AIによる既存の著作物の創作的表現に類似した生成物の出力は、通常、付随性又は軽微性等の要件を満たすとはいえず、同法47条の5は原則として適用されない⁴⁸。とりわけ、低ランク適応（LoRA）のように学習対象が少なく、許諾を得ることも容易な場合には、情報解析に係る著作権の制限規定（著作権法30条の4及び同条47条の5）を適用すべきでないと指摘されている⁴⁹。

2. 5. 小括

2. 5. 1. 機械学習と許諾の要否（解釈論）

平成30年著作権法改正による柔軟な権利制限規定の整備にあたっては、「①著作物の本来の利用には該当せず、権利者の利益を通常害さないと評価できる行為類型（第1層）」、「②著作物の本来の利用には該当せず、権利者に及び得る不利益が軽微な行為類型（第2層）」及び「③公益的政策実現のために著作物の利用の促進が期待される行為類型（第3層）」の3つが区別された⁵⁰。著作権法30条の4は第1層、同法47条の5は第2層にそれぞれ対応するものである⁵¹。そこでは、著作権の制限を正当化するにあたって、AI開発者等の営業の自由やAIシステムの社会実装による効用の増大が直接の根拠とされていたわけではない。しかし、すでに見たように、機械学習のために大規模なデータの利用が必要となる場合に、これに含まれる著作物について個別に著作権者の許諾を得なければならないとすれば、

言]。

⁴⁶ 平成30年改正解説31頁、加戸・前掲注9)419頁。

⁴⁷ たとえば、「患者の病状を踏まえて、過去の症例、治療方法、薬効等に関する様々な情報から最適な治療方法を分析し、その結果を提供するサービスにおいて、結果の提供とともに最適な治療方法と判断した根拠となる情報の一部を提供する行為」（基本的な考え方30頁以下）を検索拡張生成を用いて行うことが考えられる（座談会242頁〔上野発言〕、246頁〔今村発言〕）。

⁴⁸ 山崎・前掲注31)56頁、奥邨・前掲注18)50頁注57、座談会243頁以下〔奥邨発言、愛知発言、横山発言、前田発言〕参照。

⁴⁹ 座談会245頁以下〔奥邨発言〕。

⁵⁰ 基本的な考え方34頁、平成30年改正解説8頁、文化庁「文化審議会著作権分科会報告書」（平成29年4月）38頁以下。

⁵¹ 基本的な考え方34頁、平成30年改正解説14頁。

禁止的な費用の増大をもたらし、又は学習済みモデルの偏向を生じさせるおそれがある。これは、AI システムの開発を阻害するだけでなく、偽情報の拡散やハルシネーションの発生率の増大のリスクを高めることになりかねない。そうであってみれば、AI 開発者等に著作権者から許諾を得ることを期待できるかどうかは、これらの著作権の制限規定の解釈にあたっても考慮されるべきである。たとえば、文章から画像を生成する学習済みモデルに低ランク適応（LoRA）を組み合わせ、既存の著作物の創作的表現に類似する生成物を出力させる行為については、学習対象著作物の著作権者を特定して、その許諾を得ることは、必ずしも困難ではない。低ランク適応には数十枚程度の画像を学習させればよく、許諾を得べき著作権者は相対的に少数にとどまるものと予想されるからである。本稿では、既存の著作物の創作的表現に類似する生成物の出力を目的としてその著作物を学習の対象とする行為は、享受目的が併存し（著作権法 30 条の 4）、又は付随性や軽微性の要件を満たさないため（同法 47 条の 5）、権利制限の対象とはならないと論じたが、その結論は、こうした利害状況からも基礎づけられる⁵²。

2. 5. 2. オプトアウトの法制化（立法論）

最後に、開発・学習段階における著作物の利用に関する立法論にも簡単に触れておきたい。

第一に、機械可読な方法によるオプトアウトの法制化が考えられる。たとえば、EU デジタル単一市場指令（Directive(EU)2019/790）4 条 3 項及びこれを受けた EU 人工知能法（Regulation(EU) 2024/1689）前文 106 項及び同法 53 条 1 項(c)では、テキストマイニング及びデータマイニングのための例外又は制限について、機械可読な方法によるオプトアウトが認められている。

たしかに、機械可読な方法によってクローリング等の対象から自動的に除外できるのであれば、権利制限の対象としなくても、AI 開発者等に禁止的な費用を課すおそれは低減する。しかし、機械可読な方法によるオプトアウトを認めれば、学習の対象が恣意的に限定され、学習済みモデルに偏向をもたらすおそれがある。たとえば、フェイクニュースを意図的に拡散する者はオプトアウトを選択しないであろうから、学習の対象となる情報において偽情報が相対的に優位を占め、対話型生成 AI サービスが偽情報や誤情報を提供する確率が高まりかねない。また、学習の対象となるデータが限定されれば、ハルシネーションの生じる確率も高まることになる⁵³。

すでに見たように、開発・学習段階においては広く情報を利用できることが決定的に重要である。そのため、法制度としては、機械可読な方法によってもオプトアウトを認めるべき

⁵² 依拠性について、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(2019)54 頁、座談会 227 頁〔横山発言〕。

⁵³ 座談会 257 頁以下〔谷川発言、奥邸発言〕。なお、個人情報保護法との関連において後述するが、robots.txt やメタタグを用いる場合には、これらを記述する権限を有する者が著作権者とは限らないこと、オプトアウトの範囲を適切に限定できないこと及び理由を要しないことなどの問題があり、必ずしも著作権の正当な行使としてオプトアウトが行われるとは限らない。たとえば、著作権法によって保護されない情報であっても、ウェブサイトの管理人等によって、その利用が拒絶されることがあり得る。

ではない⁵⁴。もっとも、AI 開発者等において、任意に機械可読な方法によるオプトアウトに対応することは問題ない。学習の対象を限定するのは、事業者の経営の自由に属するからである。

第二に、契約や技術による著作権の制限規定のオーバーライドについて規律を設けることが考えられる⁵⁵。例えば、ドイツ著作権法は、契約によるオーバーライドについて、著作権の制限規定の片面的強行法規化（同法 45d 条、同法 60g 条、同法 61g 条等参照）や強行法規化（同法 55a 条 3 文、同法 69d 条 2 項、同法 69g 条 2 項等参照）と評価できる個別規定を有している。また、技術によるオーバーライドについては、同法 95b 条が、技術的保護手段を講じた権利者は、同 1 項各号に定める著作権の制限規定の受益者が著作物を利用するために必要不可欠な手段を提供しなければならないと定めている。

まず、契約によるオーバーライドについては、著作権の制限規定の適用がある場合には、その限りにおいて著作物の自由な利用が認められているのであり、私人間の契約によって著作権の行使を認めるべきではない。制限規定は著作権の限界を画定するものであり、これを私人間の合意によって変更することはできないからである。現行法の解釈としても、契約に反する情報の利用は、債務不履行責任を構成するにとどまると理解されている⁵⁶。

次に、技術によるオーバーライドについては、著作権者への経済的利益の還流を考慮して制度設計を行う必要がある。たとえば、デジタル単一市場指令 4 条 3 項及びこれを受けたドイツ著作権法 44b 条 3 項はテキストマイニング及びデータマイニングのための著作権の制限についてオプトアウトを認めているが、学術研究目的の利用を対象とする同指令 3 条及びこれを受けた同法 60d 条では認められていない。これは、営利目的のデータマイニング等について、権利者に対価を獲得する機会を与えるものといえる。そうすると、仮にドイツ著作権法 95b 条のように著作権者に著作権の制限規定を実現するための手段を提供する義務を課したとしても、少なくとも営利目的の利用については、権利者に対価を得る機会を保障する必要があるように思われる。もっとも、著作権の制限規定の技術によるオーバーライドを禁止すべきかどうかについても、日本法においては慎重な議論が必要であろう。

2. 5. 3. 小括

以上のように、本稿では、既存の著作物の創作的表現に類似する生成物を出力させることを目的とする例外的な場合を除いて、開発・学習段階における著作物の利用は著作権法 30 条の 4 第 2 号の権利制限の対象になるものと考える。学習済みモデルの用途は開かれており、生成・利用段階において生成物が既存の著作物の利用市場との衝突やその潜在的販路の阻害を生じ得るとしても、これを理由として開発・学習段階における著作物の自由な利用を制限すべきではない。たとえば、画像を学習の対象とする AI モデルが、必ずしも画像生成

⁵⁴ 機械可読な方法によるオプトアウトの問題点について、座談会 256 頁〔愛知発言〕、258 頁以下、262 頁以下〔前田発言〕、266 頁以下〔奥邨発言〕参照。

⁵⁵ 情報社会指令(2001/29/EC)6 条 4 項及びこれを受けたドイツ著作権法 95b 条について、栗田昌裕「アーキテクチャによる法の私物化と著作権制度——ドイツ法との比較法による検討」田村善之＝山根崇邦編著『知財のフロンティア第 1 巻 学際的研究の現在と未来』（勁草書房・2021）191 頁参照。

⁵⁶ 座談会 253 頁〔今村発言〕、254 頁〔上野発言〕。

システムとして用いられるわけではなく、画像認識のために用いられることもある⁵⁷。このとき、開発・学習段階における著作物の利用を一律に禁止することが適切な解決であるとは限らない。クリエイターの懸念には、既存の著作物の創作的表現に類似した生成物の出力や利用に対して著作権の行使を認めることなどによって対応すべきである。

また、著作権法による保護の対象にならないとしても、自由競争の範囲を逸脱し、営業を妨害するものであれば、営業上の利益の侵害として不法行為を構成し得る⁵⁸。たとえば、特定のクリエイターの作風等を模倣できることを売り文句として生成 AI サービスが提供され、又は生成 AI システムが販売されている場合には、営業上の利益の侵害として、不法行為法上の保護を請求することが考えられる。

3. 個人情報保護法と機械学習

3. 1. 序

個人情報保護法（以下「個情法」という。）は、著作権法のように情報を客体とする支配権を認めるものではなく、「個人の権利利益を保護する」（同法 1 条）という目的は、主として個人情報取扱事業者の義務や個人情報保護委員会（以下「個情委」という。）による監督等を通じて実現される。

個人情報取扱事業者が個人情報等の取得にあたって負うべき主な義務としては、①個人情報の不適正取得の禁止（個情法 20 条 1 項）、②要配慮個人情報の取得に係る本人同意原則（同法 20 条 2 項）及び③利用目的の通知公表義務（同法 21 条）が挙げられる。また、学習用データセットや学習済みモデルを第三者に販売等する際には、④第三者提供規制（同法 27 条以下）も問題になる。以下では、AI システムの開発・学習段階における個人情報等を含んだ情報の収集又は抽出に際して、個情法上、どのような規制を受けるかを概観する。

なお、生成 AI サービスに関連して、個情委は、令和 5 年 6 月 2 日、「生成 AI サービスの利用に関する注意喚起等」⁵⁹及び「OpenAI に対する注意喚起の概要」（以下「OpenAI 注意喚起」という）⁶⁰を公表している。このうち、OpenAI 注意喚起 1 (1)は、特定の個人情報取扱事業者に対する注意喚起であるが（個情法 147 条参照）、個人情報取扱事業者による機械学習のための情報の収集について直接に言及するものであり、現行法に関する同委員会の立場を表すものとして参考になるため、以下、同文書を必要に応じて参照する。

3. 2. 適正取得規制

個人情報とは、生存する個人に関する情報であつて、特定の個人を識別できるものをいう

⁵⁷ 座談会 258 頁〔奥邨発言〕。画像認識技術は、画像読み上げ機能、ソフトスマートロボットの制御システム、顔識別機能付カメラシステム等に幅広く用いられる。

⁵⁸ 最判平成 23 年 12 月 8 日民集 65 卷 9 号 3275 頁〔北朝鮮映画事件〕参照。

⁵⁹ 個人情報保護委員会「生成 AI サービスの利用に関する注意喚起等」（令和 5 年 6 月 2 日）（https://www.ppc.go.jp/files/pdf/230602_alert_generative_AI_service.pdf 最終閲覧日 2025 年 10 月 15 日）

⁶⁰ 個人情報保護委員会「OpenAI に対する注意喚起の概要」（令和 5 年 6 月 2 日）（https://www.ppc.go.jp/files/pdf/230602_alert_AI_utilize.pdf 最終閲覧日 2025 年 10 月 15 日）

(個情法 2 条 1 項各号)。個人情報の取得には、本人の同意は必要とされない⁶¹。ただし、「偽りその他不正の手段」による取得は禁止されている(同法 20 条 1 項)。

3. 2. 1. 「取得」

個人情報の「取得」とは、その提供を「受ける」行為をいう⁶²。個人情報を含む情報がインターネット等により公にされている場合であって、単にこれを閲覧するにすぎず、転記等を行わない場合は、「取得」にはあたらない⁶³。これに対して、クローリング等により個人情報を含む情報を収集又は抽出する行為は原則として個人情報の「取得」にあたるが⁶⁴、ただちに「不正の手段」による取得と評価されるわけではない。

3. 2. 2. 「不正の手段」

表 1. 個人情報取扱事業者が不正の手段により個人情報を取得している事例

- | |
|--|
| 事例 1) 十分な判断能力を有していない子供や障害者から、取得状況から考えて関係のない家族の収入事情などの家族の個人情報を、家族の同意なく取得する場合 |
| 事例 2) 法第 27 条第 1 項に規定する第三者提供制限違反をするよう強要して個人情報を取得する場合 |
| 事例 3) 個人情報を取得する主体や利用目的等について、意図的に虚偽の情報を示して、本人から個人情報を取得する場合 |
| 事例 4) 他の事業者に指示して不正の手段で個人情報を取得させ、当該他の事業者から個人情報を取得する場合 |
| 事例 5) 法第 27 条第 1 項に規定する第三者提供制限違反がされようとしていることを知り、又は容易に知ることができるにもかかわらず、個人情報を取得する場合 |
| 事例 6) 不正の手段で個人情報が取得されたことを知り、又は容易に知ることができるにもかかわらず、当該個人情報を取得する場合 |

(出典) 個情委「個人情報保護法ガイドライン(通則編)」3-3-1(42 頁)

個情委は、表 1 のとおり、「不正の手段により個人情報を取得している事例」を掲げている。これに即して見ると、クローリング等は、事例 1) から事例 5) までのいずれにもあらず、クローリング等を実施する事業者においては、通常、その過程で収集又は抽出される

⁶¹ 宍戸常寿監修・岡田淳ほか著『個人情報保護法』(商事法務・2024) 159 頁。

⁶² 宍戸監修・前掲注 61) 159 頁、岡村久道『個人情報保護法[第 4 版]』(商事法務・2022) 124 頁*2、石井夏生利ほか編著『個人情報保護法コンメンタール[第 2 版]』(勁草書房・2025) 302 頁〔森亮二〕参照。

⁶³ 個情委「個人情報の保護に関する法律についてのガイドライン(通則編)」(平成 28 年 11 月〔令和 7 年 6 月一部改正〕)(以下「ガイドライン(通則編)」という。)3-3-1 ※1(42 頁)。

⁶⁴ 個情委『「個人情報の保護に関する法律についてのガイドライン」に関する Q&A』(平成 29 年 2 月 16 日〔令和 7 年 7 月 1 日更新〕)(以下「Q&A」という)4-4(27 頁)。

情報が不正の手段で取得されたものであることを知り得ないため、原則として、事例6)にも該当しない。ここでいう「不正の手段」は、不適法な手段に限られず、社会通念により個別事例ごとに判断すべきであるが⁶⁵、その典型例は隠し撮りや未成年者等の不十分な判断能力につけ込む行為であり⁶⁶、取得の事実が本人に示されていないというだけの理由でクローリング等を「不正の手段」と評価することはできない⁶⁷。

もっとも、個情委は、「ダークウェブ上で掲載・取引されている個人情報に当該ダークウェブからダウンロード等により取得すること」については、「ダークウェブ上で掲載・取引されている個人情報は、掲載した者が偽りその他不正の手段により取得した個人情報である蓋然性が高く、また、掲載した者が法〔個情法〕第27条第1項に違反してこれをダークウェブ上で提供している蓋然性が高い」という理由から、個情法第20条第1項に違反するおそれがあるとしている⁶⁸。しかし、これは「掲載した者」が不正の手段により取得し、または第三者提供規制に反して提供していることを前提としており、インターネット等により公にされている情報一般に当てはまるものではない。したがって、クローリング等による個人情報の収集又は抽出は、それがダークウェブ上の情報を対象とするものでない限り、これには該当しない。

以上のように、開発・学習段階における偶発的な個人情報の取得については、個情法20条1項の適正取得規制は原則として適用されない。

3. 2. 3. robots.txt と「不正の手段」

ただし、クローリング等に用いられるプログラム(クローラー又はスクレイパー。以下「クローラー等」という。)は、通常、ウェブサイトのルートに配置された robots.txt ファイルに記述されたルールに従うように設計されている(なお、以下の記述は、メタタグについても基本的には同様に考えることができる)。このルールでは、ユーザーエージェントごとに、アクセスを許可し、又は許可しないディレクトリやファイルを指定することができる。たとえば、robots.txt に「Googlebot という名前のユーザーエージェントには、サイト全体でクロールを禁止する」というルールを記述すれば、クローラー等は原則としてこれに沿った挙動をする⁶⁹。

robots.txt に関する法制度の例として、著作権法では、所在検索サービスに係る著作権の制限規定(同法47条の5第1項1号)において、robots.txt への対応が権利制限の対象とされるための要件の一つとなっている(同法施行令7条の4第1項1号ロ、同法施行規則

⁶⁵ 岡村・前掲注62)237頁以下*2、宍戸監修・前掲注61)160頁、石井ほか編著・前掲注62)300頁以下〔森〕。

⁶⁶ 宇賀克也『新・個人情報保護法の逐条解説』(有斐閣・2021)213頁。

⁶⁷ 石井ほか編著・前掲注62)300頁以下〔森〕参照。同書は、「取得の事実を容易に示すことができるにもかかわらず、これを示さずに取得する場合には、『偽りその他不正の手段』に当たる可能性が高いと解すべき」と指摘する。

⁶⁸ Q&A4-5(27頁以下)。

⁶⁹ Google LLC「robots.txt の書き方、設定と送信」(最終更新日 UTC 2025 年 8 月 4 日)(<https://developers.google.com/search/docs/crawling-indexing/robots/create-robots-txt?hl=ja> 最終閲覧日 2025 年 10 月 15 日)。同様の機能を有するものとして、HTML で記述される robots メタタグがある。

4条の4第1号)⁷⁰。これを参考に、個人情報においても、意図的に robots.txt の記述を無視して禁止されたファイルから個人情報を含む情報を取得するように設計されたクローラーを用いる行為を「不正の手段」と評価する余地がある。そこで、以下にその適否を検討する。

たしかに、robots.txt への対応は、学習用データを収集又は抽出する事業者に期待可能な負担といえる。しかし、①個人情報法は、そもそも個人情報の取得について本人の事前同意を要求していない。その趣旨からすれば、機械可読な方法による不同意の表明によって、個人情報の取得を違法とすることを認めるべきではない。また、②robots.txt を記述する権限を有するウェブサイトの管理人が、そのサイトにおいて公開されている個人情報の本人とは限らない。さらに、③robots.txt の記述はファイルを最小単位としており、④禁止の理由も示されない。そのため、仮に個人情報法において robots.txt の遵守を義務化すれば、法の趣旨を超えて、過度に広汎なクローリング等の禁止が行われるおそれがある。たとえば、およそ個人情報を含まないファイルであっても、robots.txt にクローリング等の禁止を記述しさえすれば、その情報の収集を回避できることになる。このように、開発・学習段階において robots.txt への対応を法的義務とすれば、学習の対象が恣意的に限定され、学習済みモデルに偏向をもたらすおそれがある。そうだとすれば、robots.txt に対応しないことを直ちに「不正の手段」と評価すべきではないだろう。

以上より、現在の使用を前提とする限り、クローリング等による個人情報の取得は、robots.txt 等の記述を遵守するかどうかにかかわらず、原則としては「不正の手段」には当たらないと考える。

3. 3. 要配慮個人情報の取得規制

3. 3. 1. 機械学習のための情報収集と OpenAI 注意喚起

次に、要配慮個人情報については、現行法上、その取得に原則として本人の事前同意が必要とされている（個人情報法 20 条 2 項柱書。本人同意原則）。要配慮個人情報が本人により公開されている場合や報道機関により報道の用に供する目的で公開されている場合等にはその限りではないが（同条 7 号、同法 57 条 1 項、同法施行規則 6 条各号）、クローリング等によって偶発的に取得される要配慮個人情報は、これらの例外規定の適用があるものには限定されない⁷¹。しかし、クローリング等の際して、収集又は抽出するおそれのある要配慮個人情報のすべてについて本人の事前同意を得ることは現実的ではなく、機械学習に禁止的な費用を課すことになりかねない。このような現行法の問題に関して、OpenAI 注意喚起 1(1)は、表 2 の各事項の遵守を求めている。

⁷⁰ ただし、「情報の提供を行わないこと」が要件であり、クローラーによる情報の収集そのものを抑止するものではない。メタタグについては、同法施行規則 4 条の 4 第 2 号が同様の規定を置いている。

⁷¹ 岡田淳ほか「AI と個人情報・プライバシー」NBL1244 号(2023)76 頁・77 頁注 9、北山昇「AI の利用と個人情報保護法上の論点」自由と正義 75 巻 6 号(2024)34 頁、松尾剛行『ChatGPT と法律実務〔増補版〕』（弘文堂・2025）71 頁（以下「ChatGPT」という。）、同『生成 AI の法律実務』（弘文堂・2025）137 頁（以下「生成 AI」という。）参照。

表 2. 機械学習のために情報を収集することに関する遵守事項

- ① 収集する情報に要配慮個人情報が含まれないよう必要な取組を行うこと。
- ② 情報の収集後できる限り即時に、収集した情報に含まれ得る要配慮個人情報をできる限り減少させるための措置を講ずること。
- ③ 上記①及び②の措置を講じてもおお収集した情報に要配慮個人情報が含まれていることが発覚した場合には、できる限り即時に、かつ、学習用データセットに加工する前に、当該要配慮個人情報を削除する又は特定の個人を識別できないようにするための措置を講ずること。
- ④ 本人又は個人情報保護委員会等が、特定のサイト又は第三者から要配慮個人情報を収集しないよう要請又は指示した場合には、拒否する正当な理由がない限り、当該要請又は指示に従うこと。

(出典) 個人情報「OpenAI に対する注意喚起の概要」(令和 5 年 6 月 2 日) 1(1)

3. 3. 2. OpenAI 注意喚起の評価(解釈論)

もっとも、これらの事項をすべて遵守したとしても、これらを本人同意原則の例外として明示する規定は存在しないため⁷²、要配慮個人情報の取得が適法とされとは限らない。そのため、OpenAI 注意喚起 1(1)の意義については、これを慎重に評価する立場もある⁷³。しかし、ある種のソフトローとしては、クローリング等による偶発的な要配慮個人情報の取得をただちに違法として規制するのではなく⁷⁴、AI 開発者等に期待可能な義務を課すことに

⁷² 齊藤浩貴＝上村哲史編著『生成 AI と知財・個人情報 Q&A』(商事法務・2024) 181 頁〔館貴也〕は、「郵便物の誤配など、事業者が求めている要配慮個人情報を手にすることとなった場合でも、当該要配慮個人情報を直ちに廃棄するなど、提供を『受ける』行為がないといえる場合には、要配慮個人情報を取得していないと整理されてきたところであり……上記遵守事項は、この解釈を踏まえたものと評価できる」と指摘する。郵便物の誤配とは異なり、クローリング等は事業者が自ら行っており、OpenAI 注意喚起 1(1)は必ずしも「廃棄」までは求めているなどの相違はあるが(同②「できるかぎり減少させるための措置」、同③「削除する又は特定の個人を識別できないようにするための措置」)、「事業者が手にすることとなった」後の行為によって「取得」性を否定する点において共通しており(Q&A4-8 (28 頁以下))、同注意喚起の一部を個人情報法 20 条 2 項柱書との関係において基礎づける見解として評価に値する。

⁷³ 松尾・ChatGPT 71 頁脚注 16、同・生成 AI 138 頁、福岡真之介＝松下外『生成 AI の法的リスクと対策』(日経 BP・2023) 129 頁参照。中崎尚『生成 AI 法務・ガバナンス』(商事法務・2024) 40 頁以下は「要配慮個人情報を収集してしまう可能性を排除できない実情を踏まえて、これら 4 点を遵守していれば規制当局としては、個人情報保護法違反として摘発することはしない、という立ち位置を暗に示そうとしている」という趣旨として理解する可能性を示しつつ、「②③は『できる限り即時に』の趣旨が必ずしも明らかではないため、他の AI 開発事業者が参照するには難しい点がある」と指摘する。

⁷⁴ OpenAI 注意喚起 1(1)②は「情報の収集後」の、③は「収集した情報に要配慮個人情報が含まれていることが発覚した場合」の遵守事項であり、いずれも要配慮個人情報の取得そのものが不可避免的に生じることを前提とするものと考えられる(同②について、中崎・前掲注 73) 40 頁、同③について、齊藤＝上村編著・前掲注 72) 181 頁〔館貴也〕参

よってデータの利活用と個人の権利利益の保護との調整を図るものとして、一定の合理性を有するものといえるだろう。

3. 3. 3. 機械学習のための例外規定（立法論）

また、クローリング等によって収集された学習用データ（生データ）を用いて構築される学習用データセットや学習済みパラメータは、特定の個人との対応関係が排斥されている限り、個人情報には該当せず⁷⁵、個人の権利利益を害するおそれは相対的に低い。

そうだとすれば、立法論としては、上記4点をさらに具体化するなどの方法で、事業者に期待可能な措置をとるべき義務を課し、これを遵守する限りにおいて本人同意原則の例外を認めることが考えられる。この点については、「正当な利益のような一般的な例外事由を設けてそれに含まれると整理するアプローチや、事業者の提供するサービスのうち一定の行為類型を切り出して例外とするアプローチ」が考えられるとして、「後者の例としては、著作権法30条の4において、情報解析サービスを含め、一定の行為類型について柔軟な権利制限規定が設けられていることは参考になる」と指摘する見解がある⁷⁶。しかし、前者については、行為規範としての明確性に欠けるほか、一般条項としての例外規定を創設する必要性について慎重な検討が必要だろう。また、後者については、著作権法30条の4の立法趣旨が「著作物の本来の利用には該当せず、権利者の利益を通常害さないと評価できる行為類型」について柔軟な権利制限規定を導入することにあるところ⁷⁷、個人情報においても同様の理由から例外規定の導入を基礎づけられるかは疑問である。たとえば、学習用データとして著作物を複製する行為は、形式的には著作物の法定利用行為である「複製」に該当するが、実質的には著作物に表現された思想又は感情の享受を目的とするものではなく、その享受により効用を得ようとする者からの対価回収の機会を損なうものではないため、権利制限の対象とされている（同法30条の4第2号参照）⁷⁸。これに対して、クローリング等による偶発的な要配慮個人情報の取得については、その「本来の利用」を論じることはできず、対価回収の機会の保障のような基準もうまく適合しない。少なくとも、財産権である著作権の制限規定とは異なる視点からの検討が必要であろう。また、著作権法30条の4を参照するのであれば、事業者の提供するサービスの行為類型よりも、要配慮個人情報を取得する目的に着目して例外規定を設計すべきように思われる（本稿2.2.4.参照）。このように考えると、情報の利活用と個人の権利利益の保護とを適切に調整するためには、OpenAI 注意喚起が示した遵守事項を参考にしつつ、AI 開発者等に期待可能な行為規範を整備したうえで、これを遵守する限りにおいて、機械学習のための要配慮個人情報の取得についてAI 開発者

照)。

⁷⁵ 学習用データセットを「統計データ」(Q&A2-5 (20 頁))又は「統計情報」(個人情報保護法ガイドライン(仮名加工情報・匿名加工情報編)(平成28年11月〔令和6年12月一部改正〕30頁)に準じて考えるものとして、曾我部真裕監修「AI 開発における学習用データの利用目的と学習済みパラメータの取扱い」NBL1254号(2023)47頁以下、同49頁以下〔中井杏〕、北山・前掲注71)33頁参照。

⁷⁶ 岡田ほか・前掲注71)76頁、穴戸監修・前掲注61)737頁。

⁷⁷ 文化庁「文化審議会著作権分科会報告書」(平成29年4月)38頁以下、基本的な考え方34頁、平成30年改正解説8頁、14頁以下参照。

⁷⁸ 平成30年改正解説14頁以下、基本的な考え方7頁、11頁。

等を免責する方向での法改正が適合的と思われる。

3. 4. 利用目的規制

個人情報の取扱いにあたっては、(a)利用目的の特定義務（個情法 17 条 1 項）、(b)利用目的による制限（同法 18 条 1 項）及び (c)利用目的の通知公表義務（同法 21 条 1 項）が課されるが、クローリング等による個人情報の取得そのものにはこれらの規制は適用されない。特定の個人との対応関係が排斥されている限り、学習用データセットは個人情報に当たらず、学習用データセットへの加工を行うことを利用目的として特定する必要はないからである⁷⁹。ただし、法的義務の有無にかかわらず、クローリング等によって取得される個人情報等の利用目的及びその仕様についてはウェブサイト等において公表しておくことが望ましい⁸⁰。たとえば、robots.txt の記述方法を予め説明しておくことによって、要配慮個人情報を誤って収集するおそれを間接的に低減させられるからである（OpenAI 注意喚起 1(1)①参照）⁸¹。

3. 5. 第三者提供規制

また、特定の個人との対応関係が排斥されている限りにおいて、学習用データセットや学習済みパラメータは個人情報に当たらないため、「個人データ」（個情法 16 条 3 項）にはあらず、第三者提供規制（同法 27 条）や越境移転規制（同法 28 条）等の適用を受けない⁸²。ただし、すでに見たように、学習用データセットへの加工前に、要配慮個人情報を削除し、または特定の個人を識別できないようにするための措置を講ずることが求められていることには留意が必要である（OpenAI 注意喚起 1(1)②・③。表 2 参照）。

⁷⁹ Q&A2-5（20 頁。統計情報への加工）。岡田ほか・前掲注 71) 75 頁以下、北山・前掲注 71) 33 頁、曾我部監修・前掲注 75) 49 頁以下〔中井〕参照。プロンプトへの入力について、水井大「生成 AI と個人情報保護法」会社法務 A2Z 2024 年 7 月号 24 頁。利用目的の特定が望ましいとするものとして、佐野史明＝手島都瑠「AI と個人情報保護法」金融法務事情 2252 号(2025)55 頁。批判的な見解として、曾我部監修・前掲注 75) 52 頁〔曾我部〕参照。

⁸⁰ OpenAI 注意喚起 1(2)は、「利用者及び利用者以外の者を本人とする個人情報の利用目的について、日本語を用いて、利用者及び利用者以外の個人の双方に対して通知し又は公表すること」を遵守事項としてあげている。もっとも、これは利用者が入力する個人情報に関するものであり、クローリング等によって収集される情報を直接の対象とするものではない。

⁸¹ ただし、すでに見たように、robots.txt への対応を一律の法的義務とすべきではない。たとえば、大規模言語モデルの構築にあたって robots.txt への対応を一律に求めれば、学習済みモデルの偏向等の弊害が生じるおそれがあるからである。このことと、AI 開発者等が自主的な取組みとして robots.txt への対応を行うことは別である。AI 開発者等には、学習用データとして必要とされる情報の種類、範囲、利用目的等に応じて、要配慮個人情報の取得を抑制するために必要な取組みとして、どのような方法を用いるかを選択する自由が認められなければならない。

⁸² 野呂悠登「AI による個人情報の取扱いの注意点」Business Law Journal 2018 年 6 月号 60 頁以下、「金融審議会資金決済ワーキング・グループ報告」（2022 年 1 月 11 日）9 頁以下、曾我部監修・前掲注 75) 50 頁以下〔中井〕、北山・前掲注 71) 37 頁、実戸監修・前掲注 61) 734 頁参照。

3. 6. 小括

個情法は、個人情報等の取得を直接に規制する規定としては、同法 20 条各項を置いているのみである。同条 1 項は個人情報について「偽りその他不正の手段」による取得を一般に禁止しており、同条 2 項は要配慮個人情報の取得について原則として本人の事前同意を要件とする。

このうち、前者の適正取得規制は、クローリング等による偶発的な個人情報の取得を禁止するものと解釈すべきではない。ウェブサイトの管理者等が robots.txt やメタタグ等の記述を通してクロールを禁止する意思を表明している際に、これを無視してクローリング等を行うことを「不正の手段」と評価する余地はあるものの、個人情報の取得に本人の事前同意は不要であること、管理者等が個人情報の本人であるとは限らないこと及びその禁止が個人情報等に限定されるとは限らないことから、これを直ちに違法とすべきではない。機械学習によって構築される学習済みモデルの種類が多岐にわたることに鑑みれば⁸³、その開発・学習段階における情報の収集は可能な限り自由であるべきだろう。

同様に、後者の本人同意原則についても、機械学習のための例外規定を設けることが望ましい。現行法では、OpenAI 注意喚起 1(1)は、特定の事業者に対する指導又は助言であるものの、事業者に期待可能な義務を課し、その営業の自由と個人の権利利益との調整を図るものとして、肯定的に評価することができる。しかし、同注意喚起の掲げる各事項を遵守しても、要配慮個人情報の取得について違法性が阻却されるとは限らない。そのため、AI 開発者等にとっては、いかなる行為規範に従えばクローリング等による情報収集を適法に行えるのか必ずしも明確ではない。そこで、法令により一定の義務を課したうえで、これに従えば本人同意を要しないという免責規定を導入することが考えられる。この点について、個情委「個人情報保護法の制度的課題に対する考え方について」（令和 7 年 3 月 5 日）1 頁以下は、「統計情報等の作成にのみ利用されることが担保されていること等を条件に、本人同意なき個人データ等の第三者提供及び公開されている要配慮個人情報の取得を可能」とし、これに AI 開発等も含まれうるとの改正の方向性を示すが、事業者の法遵守を確保するための実効的な枠組みを確立することが重要であろう。

なお、機械学習のための情報収集に対する個情法の規制は、技術の発展とも密接に関連している。たとえば、技術的措置等により、(a)クローリング等による学習用データの収集・抽出時、(b)学習用データの保存時及び(c)学習用データセットへの加工時のそれぞれにおいて、（要配慮）個人情報を除外するようにフィルタリングを行い、又は個人識別性を失わせることが考えられる（OpenAI 注意喚起 1(1)①～③参照）。その技術的措置等の精度が十分に高ければ、偶発的な（要配慮）個人情報の取得そのものが例外的にしか生じなくなるだろう。

なお、EU 法では、一般データ保護規則（General Data Protection Regulation: GDPR, Regulation(EU)2016/679）5 条以下に基づき、パーソナルデータの処理（processing of personal data）が規制されている。パーソナルデータの処理は、同規則 6 条 1 項各号が列挙する法的根拠のいずれかを満たしていなければならず（適法性要件）、クローリング等によるデータ処理は、データ主体の明確な同意（explicit consent）等の要件を満たしていな

⁸³ 総務省＝経済産業省「AI 事業者ガイドライン〔第 1.1 版〕」（令和 7 年 3 月 28 日）9 頁参照。

ればならない⁸⁴。日本国と EU は相互認証を行っているため、その限りでの対応は必要であり、現に行われている⁸⁵。もっとも、日本法の立法論としては、前述の理由から、少なくとも機械学習のための情報収集について「個人情報」の取得に本人同意原則を導入する必要はないだろう。個人情報等の取得が偶発的なものであれば、取得そのものを違法評価するよりも、取得後に個人識別性を失わせるなどの措置を義務づける方が実効的であり、散在するデータの自由な利活用にも資するからである。

4. 民法上のプライバシー

最後に、民法上のプライバシー侵害との関係についても、判例法を中心に簡単に触れておきたい。

4. 1. 「宴のあと」事件判決（東京地判昭和 39 年 9 月 28 日判時 385 号 12 頁）

東京地判昭和 39 年 9 月 28 日（「宴のあと」事件）は、「プライバシーの侵害に対し法的な救済が与えられるためには、公開された内容が（イ）私生活上の事実または私生活上の事実らしく受け取られるおそれのあることがらであること、（ロ）一般人の感受性を基準にして当該私人の立場に立った場合公開を欲しないであろうと認められることがらであること、換言すれば一般人の感覚を基準として公開されることによって心理的な負担、不安を覚えるであろうと認められることがらであること、（ハ）一般の人々に未だ知られていないことがらであることを必要とし、このような公開によって当該私人が実際に不快、不安の念を覚えたことを必要とする」と判示しており、ここでいう（イ）私事性、（ロ）私秘性（秘匿性）及び（ハ）非公知性がいわゆる「プライバシー侵害の三要件」として、その後の下級審裁判例に大きな影響を与えた⁸⁶。また、これらの要件を満たす情報が、判例においても、プライ

⁸⁴ 宮下紘『EU 一般データ保護規則』（勁草書房・2018）51 頁以下、石井夏生利『EU データ保護法』（勁草書房・2020）39 頁以下参照。なお、いわゆるセンシティブデータに相当する「人種若しくは民族の出自、政治的見解、宗教的若しくは思想的信条又は労働組合への加入を明らかにするパーソナルデータの処理及び遺伝子データ、自然人を一意に識別する目的での生体データ、健康に関するデータ並びに自然人の性生活又は性的志向に関するデータの処理」は同規則 9 条 1 項により原則として禁止されており、これを行うためには同 2 項各号の要件を満たす必要がある。その一つとして、データ主体の「明確な同意」が規定されている（同 2 項(a)）。翻訳として、個情委「個人データの取扱いと関連する自然人の保護に関する、及び、そのデータの自由な移転に関する、並びに、指令 95/46/EC を廃止する欧州議会及び理事会の 2016 年 4 月 27 日の規則(EU) 2016/679（一般データ保護規則）【条文】」も参照。

⁸⁵ 2019 年 1 月 23 日、個情委は EU を「個人の権利利益を保護する上で我が国と同等の水準にあると認められる個人情報の保護に関する制度を有している外国として個人情報保護委員会規則で定めるもの」（個情法 28 条 1 項かつこ書、同法施行規則 15 条 1 項各号、平成 31 年個情委告示第 1 号）として指定し、同日、欧州委員会は日本に一般データ保護規則 45 条に基づく十分性認定を行った。

⁸⁶ 東京高決昭和 45 年 4 月 13 日高民集 23 巻 2 号 172 頁（「エロス＋虐殺」事件）、東京地判昭和 62 年 11 月 20 日民集 48 巻 2 号 218 頁（ノンフィクション「逆転」事件第一審）、東京地判平成 12 年 2 月 29 日判タ 1028 号 232 頁（「中田英寿の詩」事件第一審）、東京地判平成 16 年 11 月 10 日裁判所ウェブサイト（週刊現代キス写真事件第一審）、東京高判平成

バシーによる保護の対象となっているという評価もある⁸⁷。

4. 2. 判例におけるプライバシーによって保護される情報の要件

しかし、判例は、少なくとも明示的には、これらの情報の属性によってプライバシーによる保護の範囲を限定するという判断構造を採用していない。たとえば、最判令和4年6月24日民集76巻5号1170頁（ツイート削除請求事件）は、「個人のプライバシーに属する事実をみだりに公表されない利益は、法的保護の対象となる」と判示したうえで、上告人が旅館の女性用浴場の脱衣所に侵入したとの被疑事実で逮捕された事実を「他人にみだりに知られたくない上告人のプライバシーに属する事実」と評価し、後述する比較衡量による違法性判断を行っている。また、最判平成15年9月12日民集57巻8号973頁（江沢民主席講演会事件）は、「学籍番号、氏名、住所及び電話番号」「〔中華人民共和国国家主席の講演会に〕参加を申し込んだ学生であること」について、「秘匿されるべき必要性が必ずしも高いものではない」としつつも、「このような個人情報についても、本人が、自己が欲しない他者にはみだりにこれを開示されたくないと考えることは自然なことであり、そのことへの期待は保護されるべきものであるから、本件個人情報、上告人らのプライバシーに係る情報として法的保護の対象となる」と判示している。このように、判例は、プライバシーによる保護の対象となる情報について、「他人にみだりに知られたくない」などの基本的な性質を有することを除けば、その属性による限定を行っていない。

4. 3. プライバシーに属する事実の公表

そのうえで、判例は、「プライバシーに属する事実」の公表については、「その事実を公表されない法的利益とこれを公表する理由とを比較衡量し、前者が後者に優越する場合に不法行為が成立する」と判示しており、同様の判断基準によって差止めの可否も判断している⁸⁸。考慮事情は事案によって微妙に異なるが、たとえば、長良川事件報道訴訟判決では、少年事件の犯人であること及び経歴や交友関係等の詳細な情報を週刊誌の記事に記載して公表した行為について、①同記事が週刊誌に掲載された当時の本人の年齢や社会的地位、②当該犯罪行為の内容、③これらが公表されることによって本人のプライバシーに属する情報が伝達される範囲と被本人が被る具体的被害の程度、④同記事の目的や意義、⑤公表時の社会的状

17年5月18日判時1907号50頁（同控訴審）、東京地判平成18年3月31日判タ1209号60頁、東京地判平成19年3月26日判タ1252号305頁等多数。下級審裁判例の分類として、佃克彦『プライバシー権・肖像権の法律実務〔第3版〕』（弘文堂・2020）35頁以下。

⁸⁷ 杉原則彦「判例解説」最判解民事篇平成15年度（下）489頁、松尾剛行『最新判例にみるインターネット上のプライバシー・個人情報保護の理論と実務』（勁草書房・2017）91頁以下、同・生成AI291頁以下。

⁸⁸ 最判平成15年3月14日民集57巻3号229頁（長良川事件報道訴訟）（不法行為）。不法行為について、最判平成6年2月8日民集48巻2号149頁（ノンフィクション「逆転」事件上告審）、最判令和2年10月9日民集74巻7号1807頁。差止めについて、最決平成29年1月31日民集71巻1号63頁（グーグル検索結果削除請求事件）、前掲最判令和4年6月24日。ただし、前掲最決平成29年1月31日は「当該事実を公表されない法的利益が優越することが明らかな場合」に検索結果の削除請求を限定する。

況、⑥同記事において当該情報を公表する必要性などが考慮すべき事情として例示されている⁸⁹。

4. 4. プライバシーに係る情報の無断開示

これに対して、任意に提供された「プライバシーに係る情報」の無断開示は、本人から開示について承諾を求めることが困難であった特別の事情がない限り、「任意に提供したプライバシーに係る情報の適切な管理についての合理的な期待」を裏切るものとして不法行為を構成すると判示されている。ここで、「プライバシーに係る情報」とは、「個人識別等を行うための単純な情報であって、その限りにおいては、秘匿されるべき必要性が必ずしも高いものではない」「個人情報」をいう⁹⁰。

4. 5. 小括

以上のように、判例は、プライバシーに属する事実の公表事例では比較衡量による違法性判断を行うのに対して、任意に提供されたプライバシーに係る情報の無断開示事例では、特別の事情のない限り、原則として違法性を認めるという判断を示している。これらの判例相互の関係については争いがあるが、プライバシーに属する事実の公表に関する判例は、その対象を「個人情報」の無断開示に限定しておらず、その情報を本人が任意に提供したことも前提としていない。これに対して、プライバシーに係る情報の無断開示事例は、「情報の適切な管理についての合理的な期待」を裏切るなどの定型的に違法性を肯定すべき事情が存在する事案に関するものといえる。そうすると、任意に提供したプライバシーに係る情報の無断開示事例のように定型的な判断に適合的な事案を除けば、プライバシー侵害の違法性判断は、原則として比較衡量によるべきものと思われる。

もっとも、クロール等によって収集又は抽出された情報に「プライバシーに属する事実」や「プライバシーに係る情報」（以下、両者を併せて「プライバシー情報」という。）が含まれていたとしても、原則としてプライバシー侵害を構成しないものと思われる。というのも、インターネット等で公になっている情報をクロール等によって取得する行為は、その対象となる情報を「公表」するものではないし、クロール等の対象となったウェブサイト等がAI開発者等に情報を「開示」したと評価することも困難だからである。

また、生成・利用段階においてプライバシー情報が応答結果として出力される可能性があることを理由として、開発・学習段階におけるクロール等による情報収集を差し止めることには困難がある。将来生ずべき侵害を予防するための差し止めには、侵害を生ずる高度の蓋然性が客観的に存在することが必要であるところ⁹¹、通常、開発・学習段階では一般に侵

⁸⁹ 前掲最判平成15年3月14日。このうち、①成年者が本人である場合には、その年齢ではなく「社会的地位や影響力」が考慮の対象となることや、⑤行為時の社会的状況に「その後の変化」が付加されるなど、その後の判例によって事案に応じた展開が見られるが、考慮事情は概ね共通している（前掲最決平成29年1月31日、最判令和4年6月24日）。考慮されるべき事情をより具体的に判示した例として、前掲最判令和2年10月9日参照。

⁹⁰ 前掲最判平成15年9月12日。最判平成29年10月23日判時2351号7頁参照。

⁹¹ 物権的妨害予防請求権について、小粥太郎編『新注民法(5) 物権(2)』（有斐閣・

害を生ずる高度の蓋然性があるとはいえないからである。

ただし、①対話型生成 AI の応答結果においてプライバシー情報が出力される場合には、それが確率的な相関関係に基づいて生成されるものであっても、その真偽にかかわらず、AI 提供者による「公表」の例として、②AI システムの生成・利用段階において他人のプライバシー情報をプロンプト等として無断で入力する行為は AI 利用者による「開示」の例として、それぞれプライバシー侵害を構成し得る⁹²。このように、民法上のプライバシー侵害についても、原則として生成・利用段階における情報の公表や開示を捕捉することが、現時点での判例法に沿うものと考えられる。

5. おわりに

本稿では、AI の開発・学習段階における情報の利用について、著作権法と個人情報保護法を中心とした領域横断的考察を試みた。いうまでもなく、両法によって保護される権利利益はその性質を異にしており、一方の解釈が他方に影響を与えるべきものではない。しかし、クローリング等による機械学習のための情報収集という局面においては、著作権者の許諾を要し、又は本人の事前同意を取得しなければならないとすれば、AI 開発者等としては、データの性質を判断したうえで、権利者又は本人を特定してその許諾又は同意を得るという費用を負担せざるを得なくなる。機械学習に必要とされるデータ量が大規模であれば、この費用は事業者に期待可能な程度を超え、禁止的なものになりかねない。また、許諾や事前同意を取得できず、又は著作物若しくは要配慮個人情報等に該当する可能性のあるデータを一律に学習の対象から除外するとすれば、学習済みモデルに偏向が生じ、社会全体の厚生を低減させかねない。たとえば、すでに指摘があるように、信頼できる情報源は、取材に要した投下資本を回収する必要があるため、情報の利用を禁止する方向に動機づけられる。これに対して、偽情報の拡散を目的とするウェブサイト等は、その目的を達成するため、情報の利用に積極的に応じることが予想される。このような情報源の偏りを度外視するとしても、学習用データの規模が縮小すれば、それだけハルシネーションの発生率が増大することになる。このように見れば、法領域を問わず、AI 開発者等に対して、期待可能性がないにもかかわらず、すべてのデータについて情報の性質を判断したうえで特定の者の意思を確認するように一律に義務づける解釈は望ましくないといえるだろう。その限りにおいて、著作権者の財産権保障や本人の人格的自律の権利は、情報の自由や営業の自由などとの調整を必要とするものである。

ただし、AI に関する急速な技術の発展と社会実装の進展に鑑みれば、この結論は暫定的なものといわざるを得ない。たとえば、本稿では、著作権法による保護の対象に画風等を追加すべきではないと考えるが、生成 AI サービスの発展の方向によっては、機械学習における著作物の利用についても著作権者に経済的参加を認めるために、何らかの権利を創設すべきことになるかもしれない。また、クローリング等の手段を用いた偶発的な要配慮個人情報の取得については、技術的な措置が十分な精度を有する場合には、「取得」そのものが例

2020) 294 頁〔金子敬明〕(占有保全の訴え)、松岡久和『物権法』(成文堂・2017) 33 頁、安永正明『講義 物権・担保物権法』(有斐閣・2021) 17 頁、山野目章夫『民法概論 2 物権法』(有斐閣・2022) 161 頁等。

⁹² 栗田昌裕「AI と個人情報」現代消費者法 68 号(2025)20 頁以下。

外的にしか生じなくなるかもしれない。

なお、民法上のプライバシーに関する判例には、現在のところ、AIの開発・学習段階におけるプライバシー情報の取扱いに直接に言及するものはない。もっとも、現在のところ、侵害行為の態様としては、プライバシー情報の公表又は無断開示が主として想定されており、その限りにおいて、クローリング等がプライバシー侵害と評価されるのは例外的な場合に限られるだろう。また、AIの生成・利用段階も含めて、プライバシー侵害が問題となり得る事例であっても、定型的な違法性判断に適合的な事例を除けば、原則としてプライバシー情報を利用されない法的利益と利用する理由との比較衡量によって違法性判断が行われるべきことになる。その点では、民法上のプライバシーに関する判例法は、発展の余地を残している。

以上のように、機械学習のための情報収集については、AI開発者等に期待することのできない費用を課すことにより、その開発を妨げ、又はその社会実装に際して必要な精度を保てなくなることがないように制度設計を行う必要がある。また、AI開発者等に個別の許諾や事前同意の取得等を期待できる場合には、個人の自律を支援し、その実質的な自己決定を担保することが求められる。いずれにせよ、AIと情報の利用との関係を論じる際には、開発・学習段階、入力段階又は出力・利用段階等の場面を分けたうえで、分析的にその利害状況を検討しなければならない。本稿では、開発・学習段階における情報の利用の一部を論じるにとどまった。その他の場面については、他稿を期したい。

(脱稿日 2025 年 9 月 13 日)

参考文献

本文中に掲げたもの。

- ※ 本稿の校了後、統計情報の作成等が目的の場合は事前の本人同意なしで要配慮個人情報の取得を可能としつつ、悪質な違反に対する課徴金制度を導入するなどの個人情報保護法改正案が公表されたとの報道に接した。これに際して本文の一部に修正を加えたが、改正案の評価と検討については他稿を期したい。