

生成AI時代の著作権リスク



STORIA

2026.2.5

STORIA法律事務所
弁護士 柿沼太一

【自己紹介】



- ▼ 1997年京都大学卒業
- ▼ 2000年4月弁護士登録
- ▼ 2015年STORIA法律事務所設立
- ▼ 2021年神戸大学大学院TLP（知的財産法）修了・博士号（法学）取得
- ▼ AI、データ、ヘルスケア、知的財産、スタートアップ法務を主として取り扱う
- ▼ 経産省の「AI・契約ガイドライン」検討委員会委員（2018）
- ▼ 経済産業省「オープンイノベーションを促進するための技術分野別契約ガイドラインに関する調査研究」委員会事務局(2021～)
- ▼ 日本データベース学会理事
- ▼ 日本ディープラーニング協会理事

■ 目次

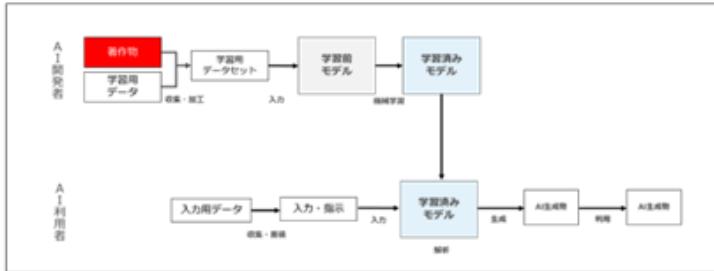
- 1 AI開発者・提供者の著作権責任が問題となるパターン
- 2 学習済みモデル生成のための既存著作物の利用（パターン1）
- 3 学習用データセットの公開（パターン2）
- 4 学習済みモデルの公開（パターン3）
- 5 RAGのための既存著作物の収集・蓄積（パターン4）
- 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

■ 目次

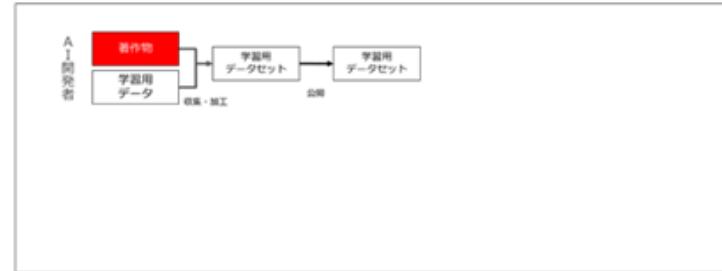
-
- 1 AI開発者・提供者の著作権責任が問題となるパターン
 - 2 学習済みモデル生成のための既存著作物の利用（パターン1）
 - 3 学習用データセットの公開（パターン2）
 - 4 学習済みモデルの公開（パターン3）
 - 5 RAGのための既存著作物の収集・蓄積（パターン4）
 - 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

1 AI開発者・提供者の著作権責任が問題となるパターン

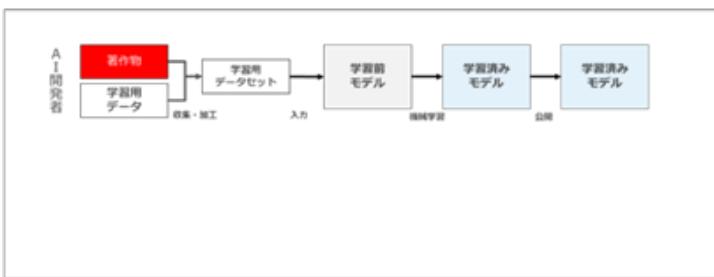
▼ パターン1



▼ パターン2



▼ パターン3



▼ パターン1

AI（学習済みモデル）生成のために既存著作物を利用する行為

▼ パターン2

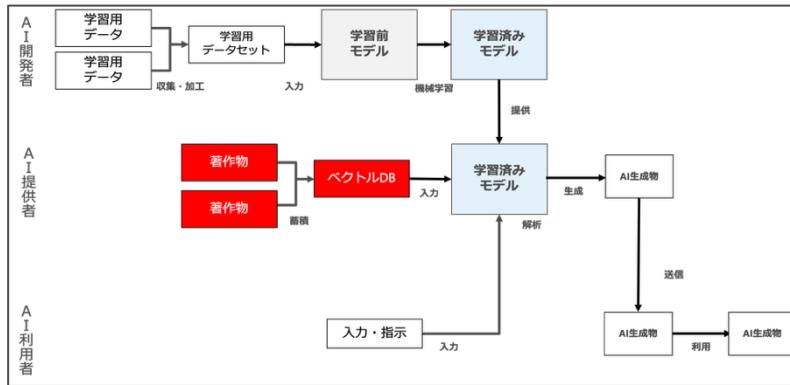
学習用データセットを公開・譲渡する行為

▼ パターン3

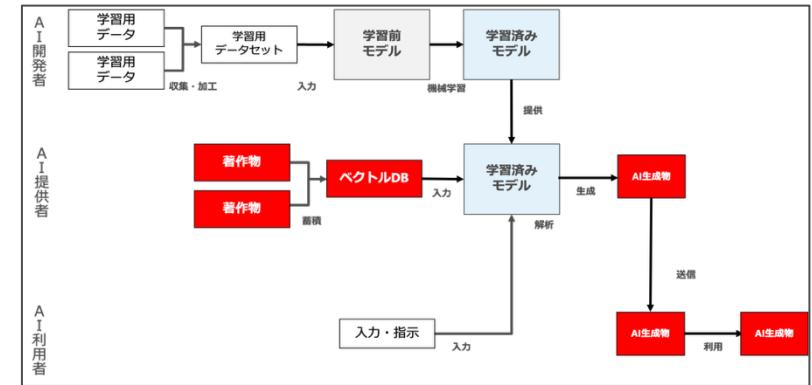
学習済みモデルを公開・譲渡する行為

1 AI開発者・提供者の著作権責任が問題となるパターン

▼ パターン4



▼ パターン5



▼ パターン4

AI提供者がRAGシステムの構築のために既存著作物を収集・蓄積する行為

▼ パターン5

AI提供者がRAGシステムの構築のために既存著作物を収集・蓄積する行為 + 蓄積した既存著作物と同一・類似のAI生成物を生成・送信する行為

1 AI開発者・提供者の責任が問題となるパターン

▼ 基礎資料

AIと著作権に関する考え方について

令和6年3月15日

文化審議会著作権分科会法制度小委員会

AI時代の知的財産権検討会
中間とりまとめ

2024年5月

AI時代の知的財産権検討会

AI事業者ガイドライン

(第1.0版)

令和6年4月19日

総務省 経済産業省

文化庁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（2024年3月15日）」

内閣府知的財産戦略推進事務局AI時代の知的財産権検討会「中間とりまとめ」（2024年5月28日）

経済産業省・総務省「AI事業者ガイドライン（第1.0版）」（2024年4月19日）

▼ AIと著作権についての基本的な考え方がまとめた資料

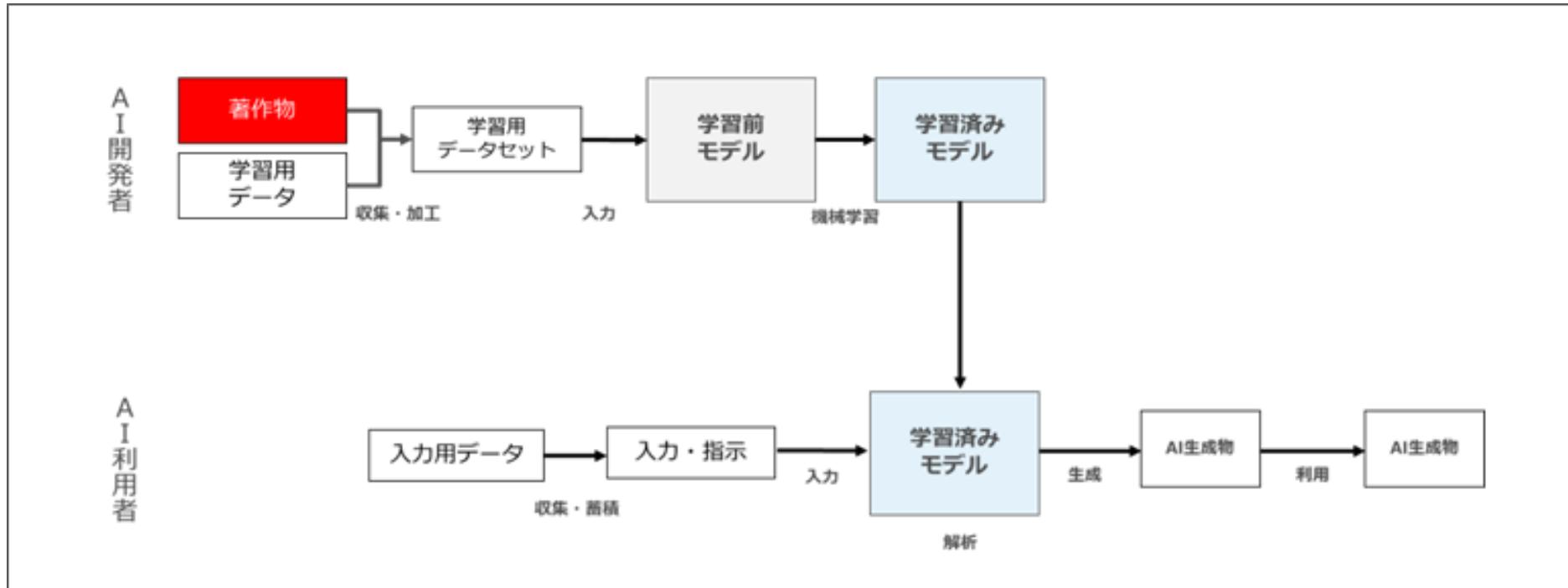
▼ AIと著作権以外の知的財産権（特許権等やパブリティ権・肖像権等）についての基本的な考え方がまとめた資料

▼ AI開発者・AI提供者・AI利用者ごとに留意すべき点（必ずしも法的な留意点には限られない）についてまとめた資料

■ 目次

-
- 1 AI開発者・提供者の著作権責任が問題となるパターン
 - 2 学習済みモデル生成のための既存著作物の利用（パターン1）
 - 3 学習用データセットの公開（パターン2）
 - 4 学習済みモデルの公開（パターン3）
 - 5 RAGのための既存著作物の収集・蓄積（パターン4）
 - 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

2 学習済みモデル生成のための既存著作物の利用（パターン1）



2 学習済みモデル生成のための既存著作物の利用（パターン1）

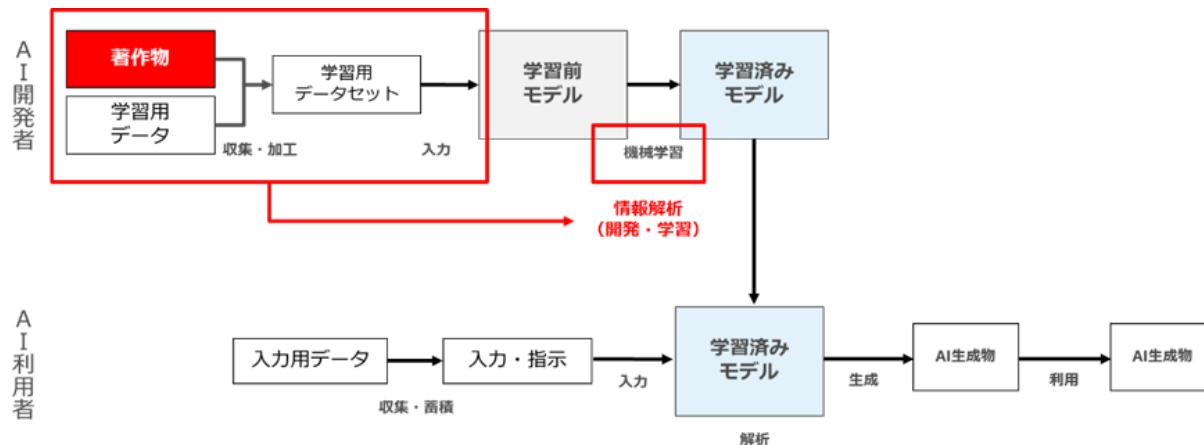
▼ 原則

法30条の4第2号により適法

（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）

第三十条の四 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。ただし、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は、この限りでない。

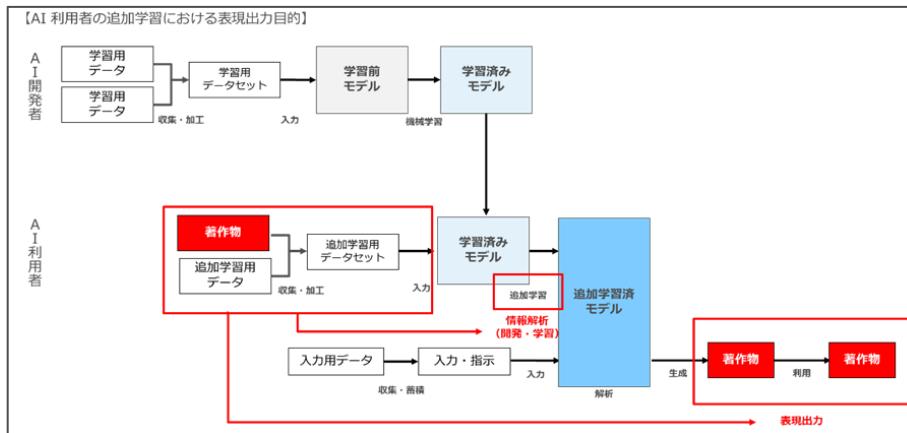
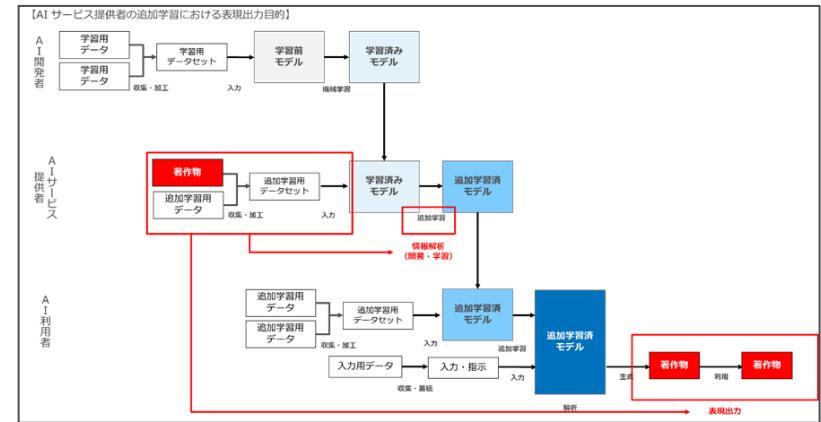
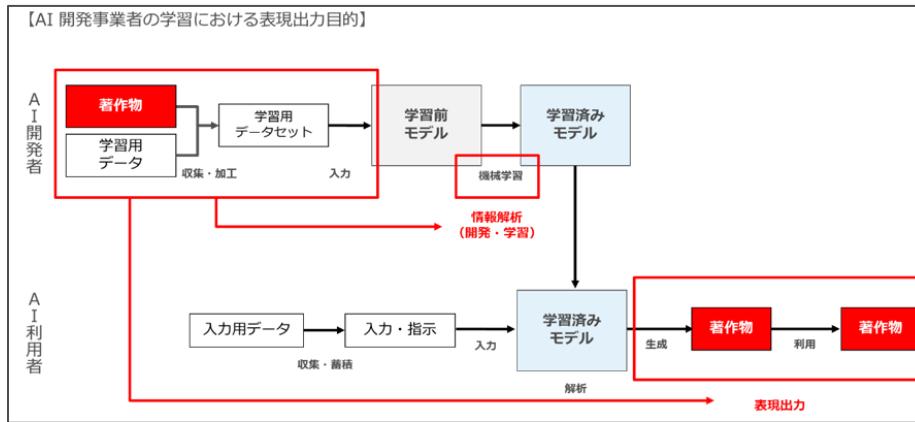
- 一 著作物の録音、録画その他の利用に係る技術の開発又は実用化のための試験の用に供する場合
- 二 **情報解析**（多数の著作物その他の大量の情報から、当該情報を構成する言語、音、影像その他の要素に係る情報を抽出し、比較、分類その他の解析を行うことをいう。第四十七条の五第一項第二号において同じ。）の用に供する場合
- 三 前二号に掲げる場合のほか、著作物の表現についての人の知覚による認識を伴うことなく当該著作物を電子計算機による情報処理の過程における利用その他の利用（プログラムの著作物にあつては、当該著作物の電子計算機における実行を除く。）に供する場合



2 学習済みモデル生成のための既存著作物の利用（パターン1）

▼ 例外

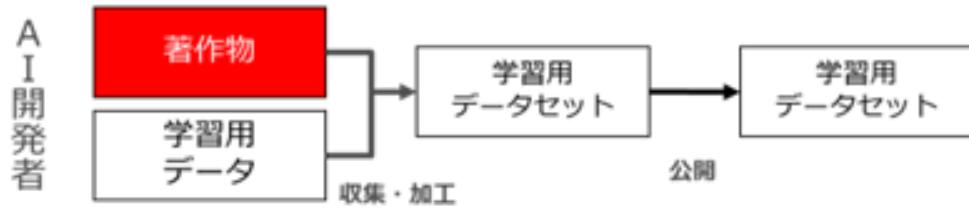
条文上は、当該既存著作物利用行為に「30条の4第2号が適用されない」か「30条の4柱書但書が適用される」のどちらかに該当する場合は、他の権利制限規定が適用されない限り著作権侵害となる。



■ 目次

-
- 1 AI開発者・提供者の著作権責任が問題となるパターン
 - 2 学習済みモデル生成のための既存著作物の利用（パターン1）
 - 3 学習用データセットの公開（パターン2）
 - 4 学習済みモデルの公開（パターン3）
 - 5 RAGのための既存著作物の収集・蓄積（パターン4）
 - 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

3 学習用データセットの公開（パターン2）



3 学習用データセットの公開（パターン2）

事例：Web上のデータやDSを大規模に収集して学習用DS及び生成AIモデルを構築し、学習用DSとAIモデルを公開したところ、学習用DSは収集対象となったデータを加工したものが含まれていることから著作権侵害になるのではないかという指摘があった。どのように考えればよいか。

▼ 原則

学習用DSの中には、収集対象になったデータ（著作物）がそのままの形式で、あるいは若干加工した形で含まれている

→当該DSの公開や提供は、当該DSの中に含まれている著作物やその加工物（二次的著作物）の公衆送信行為や譲渡行為に該当し、当該行為は、原則として元のデータ（著作物）の著作権者の承諾が必要

→しかし、著作権法30条の4は「情報解析（略）の用に供する場合」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」と定めているため、**他人のために学習用DSを作成することや、同DSの公開も同条により適法。**

【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関する基本的な考え方（著作権法30条の4、47条の4及び47条の5関係）】

問 11 人工知能の開発に関し、人工知能が学習するためのデータの収集行為、人工知能の開発を行う第三者への学習用データの提供行為は、それぞれ権利制限の対象となるか。

【回答】（前略）

また、収集した学習用データを第三者に提供する行為についても、当該学習用データの利用が人工知能の開発という目的に限定されている限りは、「著作物に表現された思想または感情を享受」することを目的としない著作物の利用に該当し、法30条の4による権利制限の対象となるものと考えられる。

■ 目次

-
- 1 AI開発者・提供者の著作権責任が問題となるパターン
 - 2 学習済みモデル生成のための既存著作物の利用（パターン1）
 - 3 学習用データセットの公開（パターン2）
 - 4 学習済みモデルの公開（パターン3）
 - 5 RAGのための既存著作物の収集・蓄積（パターン4）
 - 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

4 学習済みモデルの公開（パターン3）



4 学習済みモデルの公開（パターン3）

▼ 原則

学習用データセットと異なり、学習済みモデルの中に、学習に用いられたデータ（著作物）の創作的表現が残っていることは通常ない

→したがって、学習済みモデルは学習に用いられた著作物の二次的著作物とは評価できず、学習済みモデルの公開・譲渡行為そのものが著作権侵害に該当することはない

▼ 例外

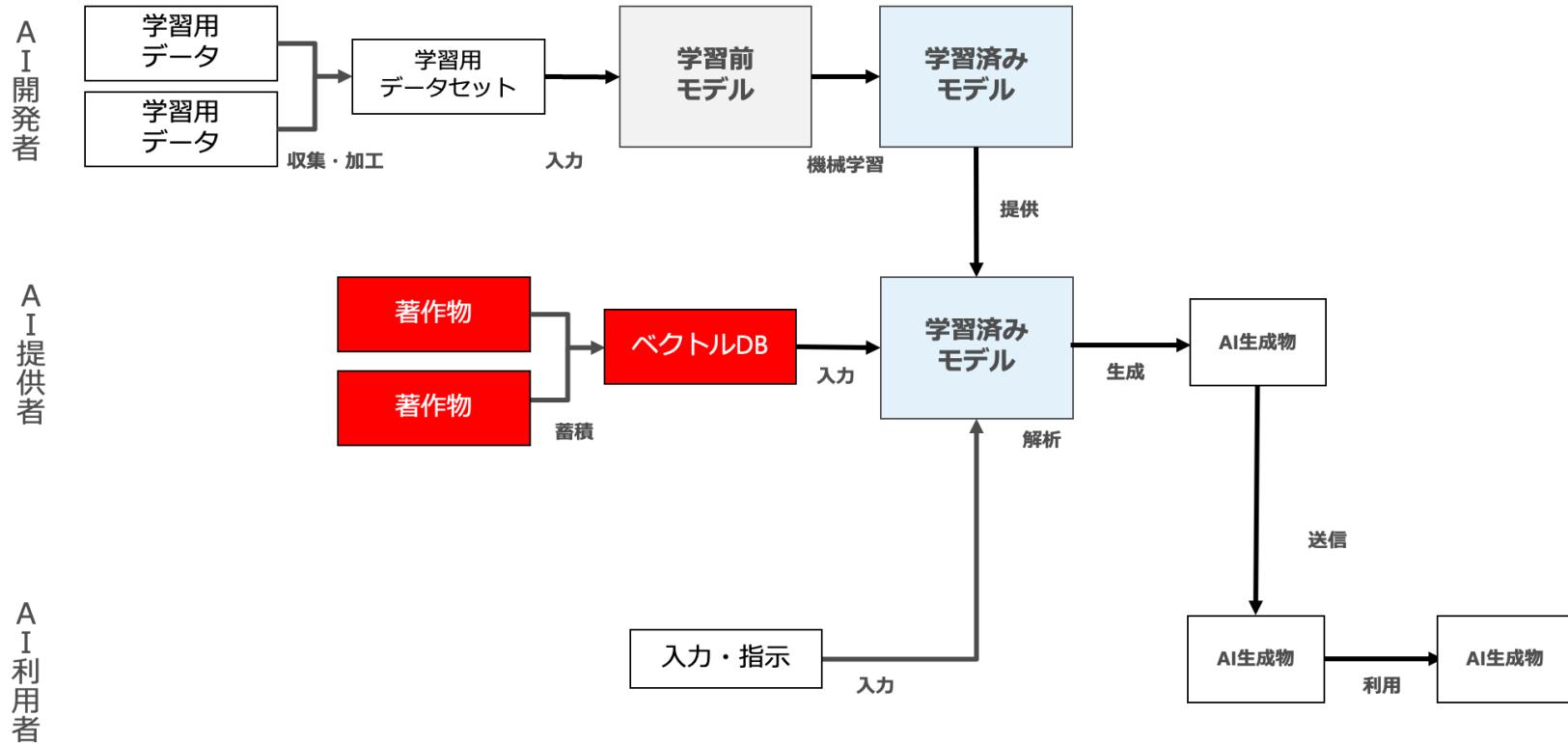
表現出力目的で学習するなど特殊な学習が行われることによって、**当該学習済モデルが、学習データである著作物と類似性のある生成物を高確率で生成する状態にある等の場合**は、当該モデルが法的には、当該学習済モデルが学習データである著作物の複製物であると評価される場合も考えられる（考え方 P 30）

→その場合は、当該学習済みモデルの譲渡は著作権侵害に該当することとなり、該著作権侵害に対する権利者の措置請求（112条2項）として、当該学習済みモデルの廃棄請求が認められる可能性がある。

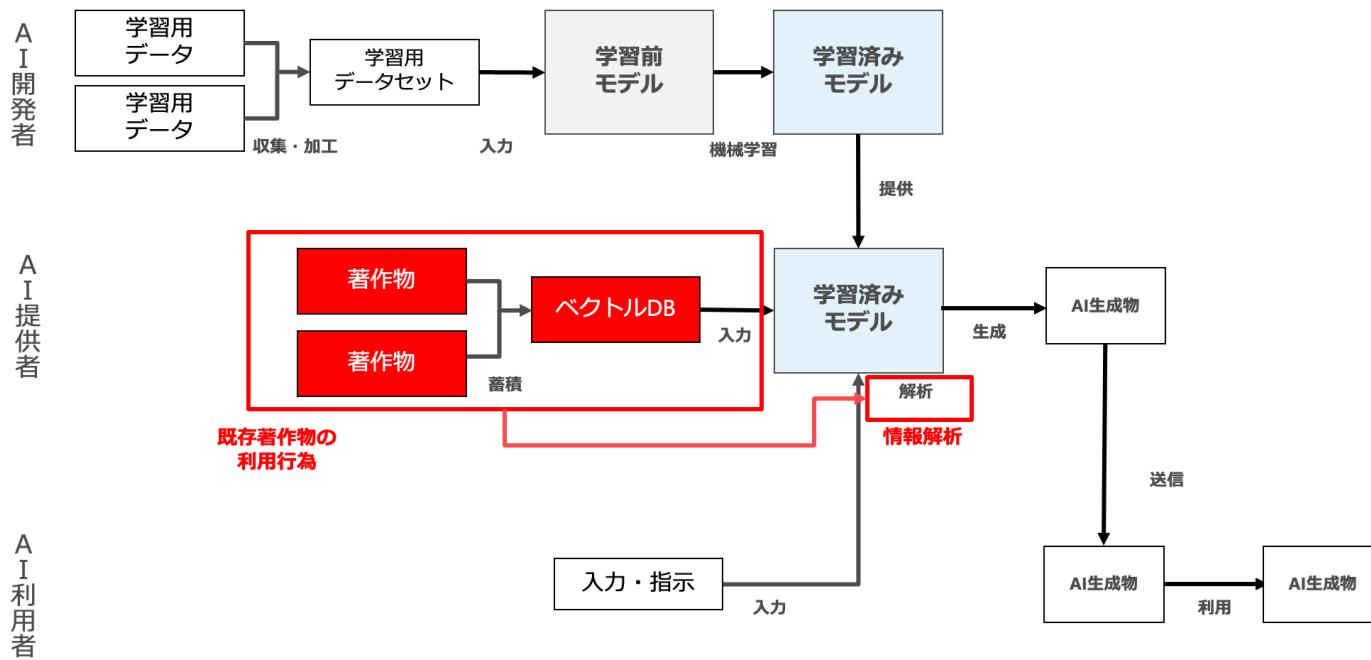
■ 目次

-
- 1 AI開発者・提供者の著作権責任が問題となるパターン
 - 2 学習済みモデル生成のための既存著作物の利用（パターン1）
 - 3 学習用データセットの公開（パターン2）
 - 4 学習済みモデルの公開（パターン3）
 - 5 RAGのための既存著作物の収集・蓄積（パターン4）
 - 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

5 RAGのための既存著作物の収集・蓄積（パターン4）



5 RAGのための既存著作物の収集・蓄積（パターン4）



- 生成AI内における「入力著作物のモデル内の解析行為」
 - 「情報解析（30条の4第2号）」あるいは同3号に該当
 - 当該「情報解析」のための既存著作物の蓄積・入力は「情報解析のために必要な行為」に該当し、30条の4第2号により**原則として適法**
 - ただし、パターン1で説明した例外（学習目的による制限及び学習対象による制限）は同様に問題となる。

5 RAGのための既存著作物の収集・蓄積（パターン4）

▼「考え方」21頁～

ウ 検索拡張生成（RAG）等について

○ 検索拡張生成（RAG）その他の、生成AIによって著作物を含む対象データを検索し、その結果の要約等を行って回答を生成する手法（以下「RAG等」という。）については、これを実装しようとする場合、開発・学習段階において、生成AI自体の開発に伴う学習のための著作物の複製等のほかに、既存のデータベースやインターネット上に掲載されたデータに含まれる著作物の内容をベクトルに変換したデータベースを作成する等の行為に伴う著作物の複製等が生じ得る（上記ア（ウ）・複製⑤）。

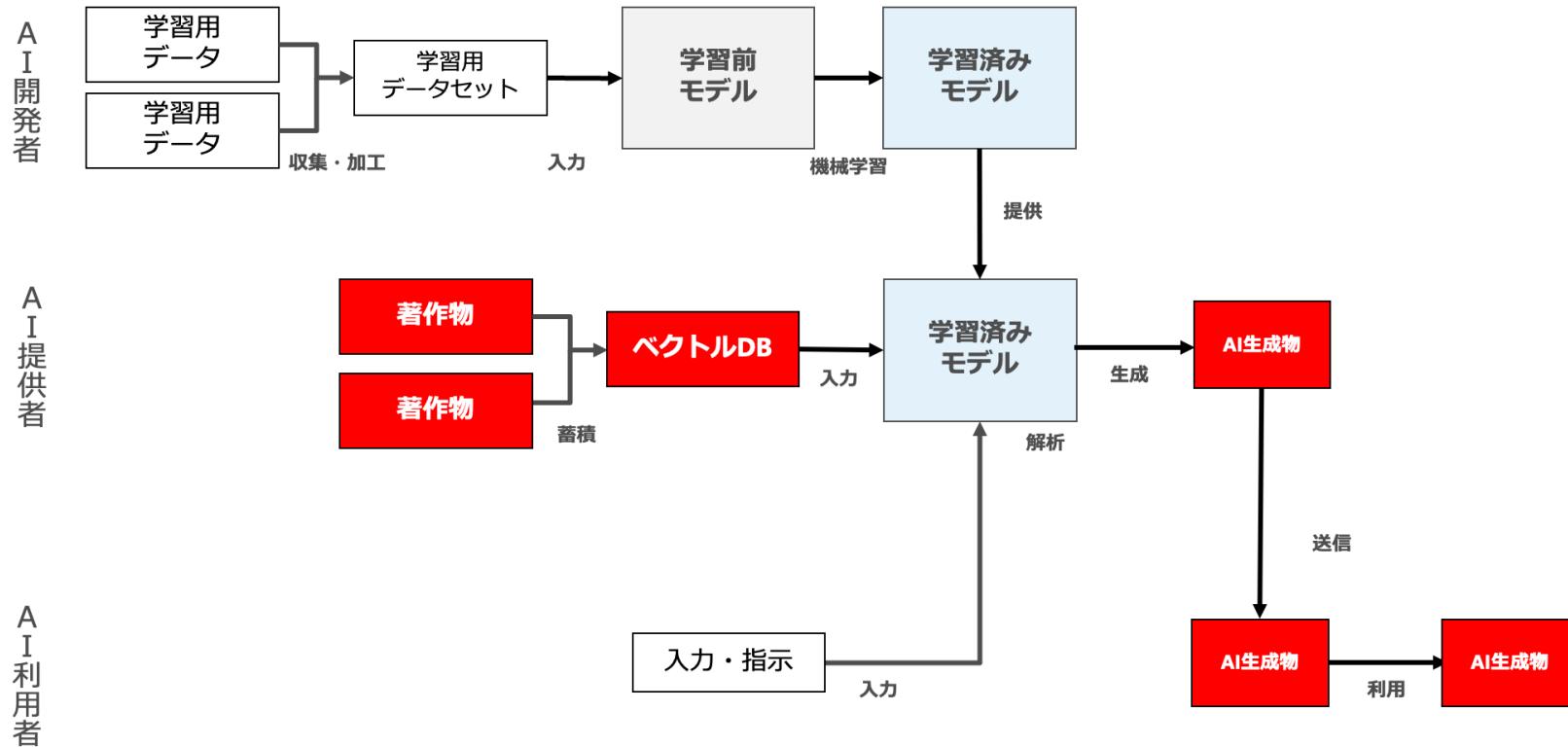
このような場合、既存のデータベースやインターネット上に掲載されたデータが著作物を含まないものであれば著作権法上の問題は生じない。また、既存のデータベースやインターネット上に掲載されたデータに著作物が含まれる場合でも、RAG等に用いられるデータベースを作成する等の行為に伴う著作物の複製等が、回答の生成に際して、当該データベースの作成に用いられた既存の著作物の創作的表現を出力することを目的としないものである場合は、当該複製等について、非享受目的の利用行為として法第30条の4が適用され得ると考えられる。

他方、既存のデータベースやインターネット上に掲載されたデータに著作物が含まれる場合であって、著作物の内容をベクトルに変換したデータベースの作成等に伴う著作物の複製等が、生成に際して、当該複製等に用いられた著作物の創作的表現の全部又は一部を出力することを目的としたものである場合には、当該複製等は、非享受目的の利用行為とはいえず、法第30条の4は適用されないと考えられる。

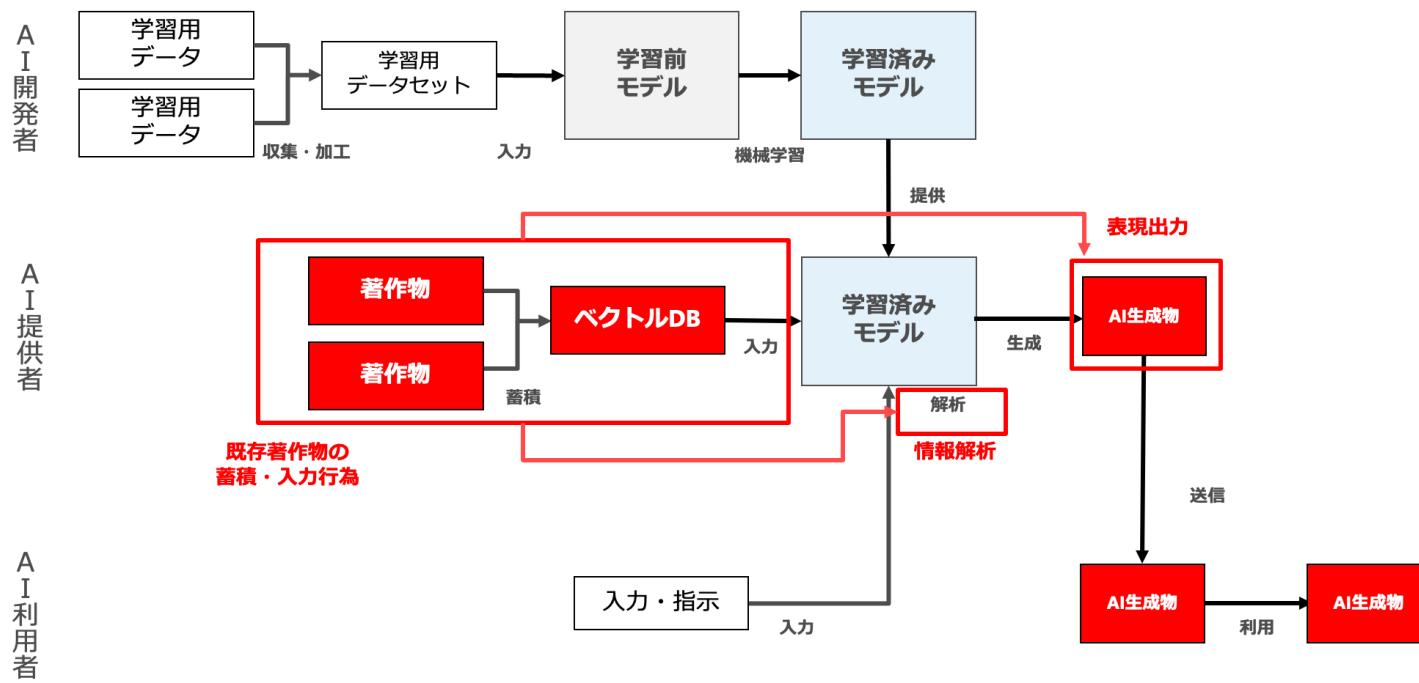
■ 目次

-
- 1 AI開発者・提供者の著作権責任が問題となるパターン
 - 2 学習済みモデル生成のための既存著作物の利用（パターン1）
 - 3 学習用データセットの公開（パターン2）
 - 4 学習済みモデルの公開（パターン3）
 - 5 RAGのための既存著作物の収集・蓄積（パターン4）
 - 6 RAGのための既存著作物の収集・蓄積+類似物生成（パターン5）

6 RAGのための既存著作物の収集・蓄積 + 類似物生成（パターン5）



6 RAGのための既存著作物の収集・蓄積 + 類似物生成（パターン5）



① 既存著作物の蓄積行為

類似AI生生物の生成が頻発するのであれば表現出力目的ありとして30条の4の適用なし。

→他の権利制限規定の適用がなければ著作権侵害

② 既存著作物に類似するAI生成物の生成・送信行為

30条の4の適用なし

→他の権利制限規定の適用がなければ著作権侵害

