

生成AIと著作権

生成AIによって生じる著作権関連の課題と対処

佐渡 秀治 Open Source Guy (2023.9.19)

本稿は2023年6月時点における国内外の法制、議論を基になるべく総和的にまとめたものであり、以降の改変は必要と思われる程度において逐次行っている。AI法制は日々目まぐるしく変化をしている世界でありながら、未だに法解釈の指針となる判例が絶対的に不足しており、本稿の内容が急速に過去のものとなる可能性に注意を要する。

AIの進化

- AI (Artificial Intelligence, 人工知能)とは、
 - 「これまで人間にしかできなかった知的な行為（認識、推論、言語運用、創造など）を、どのような手順（アルゴリズム）とどのようなデータ（事前情報や知識）を準備すれば、それを機械的に実行できるか」（日本大百科全書 ニッポニカ）
- 創作を開始したAI
 - 識別AIのような人間の作業の補助から徐々にAI自身が創作する方向へ



作曲AI「Iamus」：スペイン マラガ大学

- 8分で自律的に楽曲を作成
- MP3や楽譜などの形式で出力
- 実際に作曲された楽曲をオーケストラが演奏し、それを収録したCDや音源の販売もされている



「The Next Rembrandt」：

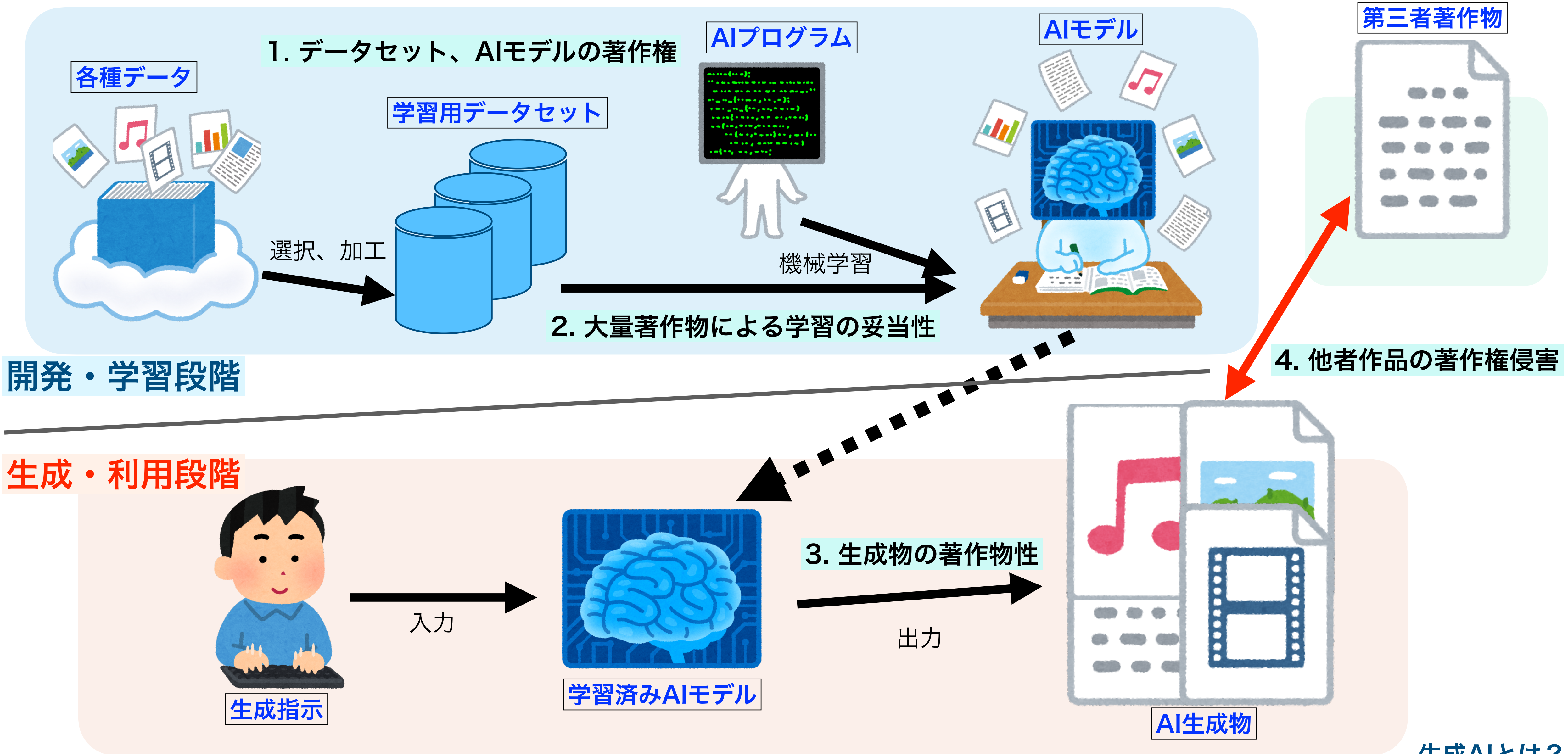
- ING、Walter Thompson、Microsoft
- レンブラントの全作品データを利用
- 色、模様、画風その他の手法の理解のために設計された深層学習アルゴリズムと顔認識技術

生成AIに関する諸問題

- 生成的AI
 - 基盤モデル上に構築したチャットシステムやアートシステム
 - 大規模言語モデル：GPT-n、拡散モデル
 - ChatGPT、Bard、Stable Diffusion、DALL-E
- 進行中の諸問題
 - プライバシー保護
 - 誤情報誘導等の悪用
 - セキュリティ、情報流出リスク
 - 社会倫理、政治的中立性の問題
 - Hallucination(幻覚)
- 著作権
 - 最も根幹的な問題として一般的に広く認識されている
 - **大量の著作物を収集して学習し、それによって開発されたモデルによって創作物を生み出す**

利用者側の意識や規約の策定で
ある程度緩和可能

生成AIにおける著作権



1. データセット、AIモデルの著作権

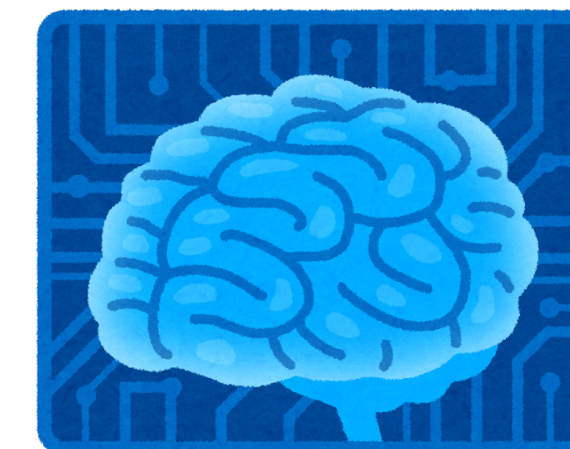
データセットの著作物性



- データは著作物ではない
 - 思想や感情を創作的に表現したものではない
- 学習用データセット(データの集合)ではどうか？
 - 🇯🇵 データベース著作物「**情報の選択**又は**体系的な構成**によつて創作性を有するもの」
 - 大量データを網羅的に収集したデータ集合体 → 「情報の選択」ではない
 - 大量のデータを蓄積しただけ → 「体系的な構成」ではない
 - 🇺🇸 データ集合に対する特別な規定はない
 - データ構造、選択、配置に必要なレベルの**創造性と独創性**を示していれば保護
 - 🇪🇺 EUデータベース指令により知的創作物に該当しないデータベースに特別な権利
 - 15年間の独占的な複製、翻案、改変等の権利保護
 - 機械的に収集されたデータベースには適用されないという解釈？
 - 基本的にはデータセット自身には知的財産権保護はほぼ及ばないと考えられる
- Common Crawl、LAION、The Pile、Wikipedia、GitHub . . .
 - 単なるデータだがクロール先には様々な**著作権が発生しライセンス付与された著作物**も含む

AIモデルの著作権

- 学習済みAIモデルは学習済みパラメータと推論プログラムから構成
 - 学習済みパラメータ
 - 学習データセットでの機械学習の結果で得られたパラメータ、ウェイトのこと
 - 外形的には機械学習によって得られた単なる数値データ
 - 著作権は認められない 🇯🇵🇪🇺🇺🇸
 - 自動的に学習され人間の関与もないので創作性は認められない
 - 情報の選択や体系的な構成に創作性がなく、データベースの著作物としての保護も難しい
 - 推論プログラム
 - 学習済みパラメータを適用して入力に対して一定の結果が出力可能にするプログラム
 - 一般のプログラムと同様に著作権が認められると考えられる 🇯🇵🇪🇺🇺🇸
 - 両者一体化したAIモデル全体の著作物性は？
 - 基本的にはプログラム部分だけの著作権が認められる
 - 使用許諾の契約等でカバーすることになる



(参考) AI・データの利用に関する契約ガイドライン (経済産業省)

https://www.meti.go.jp/policy/mono_info_service/connected_industries/sharing_and_utilization/20180615001-1.pdf

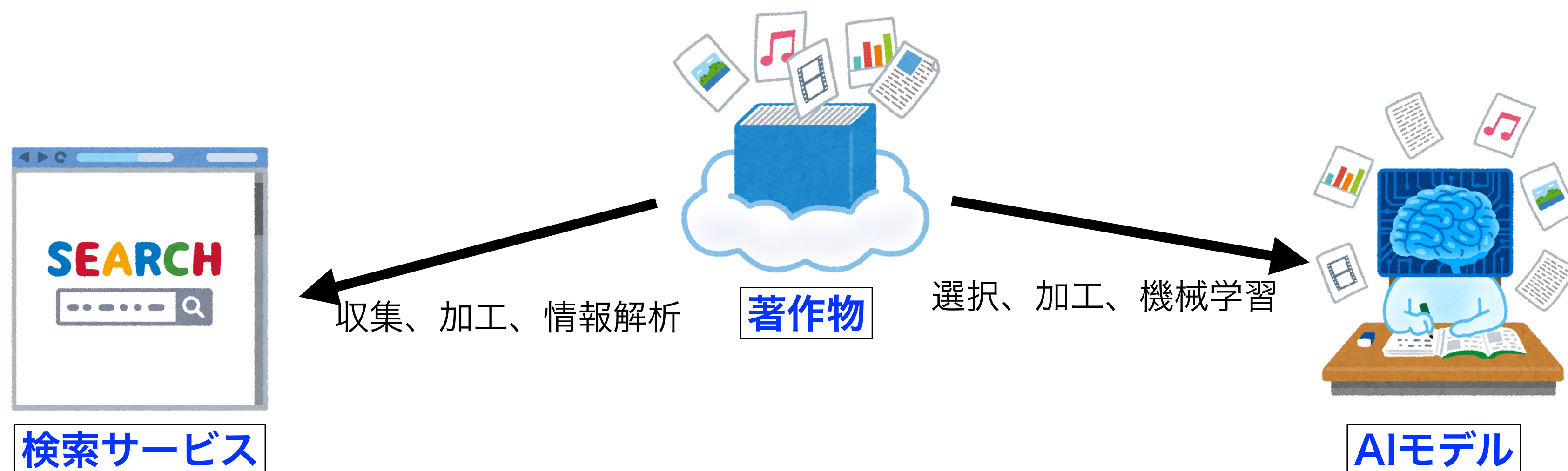
オープンに頒布されるAIモデル

- 学習済みモデルは一般に広く頒布されるものも多い
 - オープンソースライセンス、クリエイティブコモンズ、OpenRAIL-M
 - 開発、利用の輪を大きくし、競争を促進するにはどれも有効
 - OpenRAIL-Mにはモデルの使用に対しての倫理的な条項が存在
 - AIモデルに関して、
 - 著作権ライセンス(許諾)は厳密には有効ではない
 - プログラム部分はライセンスで扱われるが、パラメータ部分には適用されないと考えられる
 - AI開発コミュニティ内ではライセンスが大体は機能しているが、今後は分からない
 - ライセンスを無視して良いという意味ではない！
 - オープンソースと詐称しても良いということでもない！
 - オープンソースAIの定義
 - 現在、Open Source Initiativeで定義策定プロセスが進行中
 - データに著作物性がないこととライセンスモデルが併立するかはよく分からない

2. 大量著作物による学習の妥当性

AIモデル開発時の権利制限規定

- AIモデルの開発には大量の(第三者)著作物もデータとして投入
 - 事実上、学習データセットは大勢の著作権者による**著作物の集合体**
 - AIモデルの学習時には著作物の集合から**機械的に情報を抽出し、解析する** *
 - 大量の著作物の**結合、複製、改変等の利用行為**が行われる
 - インターネットの検索サービス等の情報解析サービスでも事前の著作物を利用した解析が必要
 - デジタル技術の台頭と共に著作権法上での柔軟な権利の制限規定が求められてきた
 - 新しい利用法の出現と共に複製等の独占権が及ぶかその都度検討
 - 本来的利用(観る、聴く、読む…)から外れ、権利者の利益を害さない利用を明確に権利制限へ



注：以降は特段の説明がなければ大規模言語モデル、拡散モデル等には学習対象著作物の「表現」が残っていないことを前提とする。

2009年改正旧著作権法第47条の7

• 2009年改正著作権法

- 情報検索サービスに必要な行為は、著作権者の 許諾を得なくても可能とすることを**明確化** (合法化ではない)
- キャッシュサーバーやバックアップサーバー等における情報の蓄積は、著作権侵害とならないことを**明確化**

• 旧第47条の7 (情報解析研究のための複製)

- 計算機による情報解析を目的とする場合には、必要と認められる限度で著作物の複製と翻案が可能に
 - この時点で諸外国よりも著作物利用の機械学習が行いやすい環境となったと考えられる
 - おそらく、現在のAIモデル開発においても条文は機能する
- ただし、情報解析の対象となる著作物の「表現」(もしくは本質的特徴)の出力の可否やその量の許容が不明瞭だった

2009(平成21)年改正著作権法 第四十七条の七

著作物は、電子計算機による情報解析（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、映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**を行うことをいう。以下この条において同じ。）を行うことを**目的とする場合**には、必要と認められる限度において、**記録媒体への記録又は翻案**（これにより創作した二次的著作物の記録を含む。）を行うことができる。(略)

柔軟な権利制限規定 (再掲)

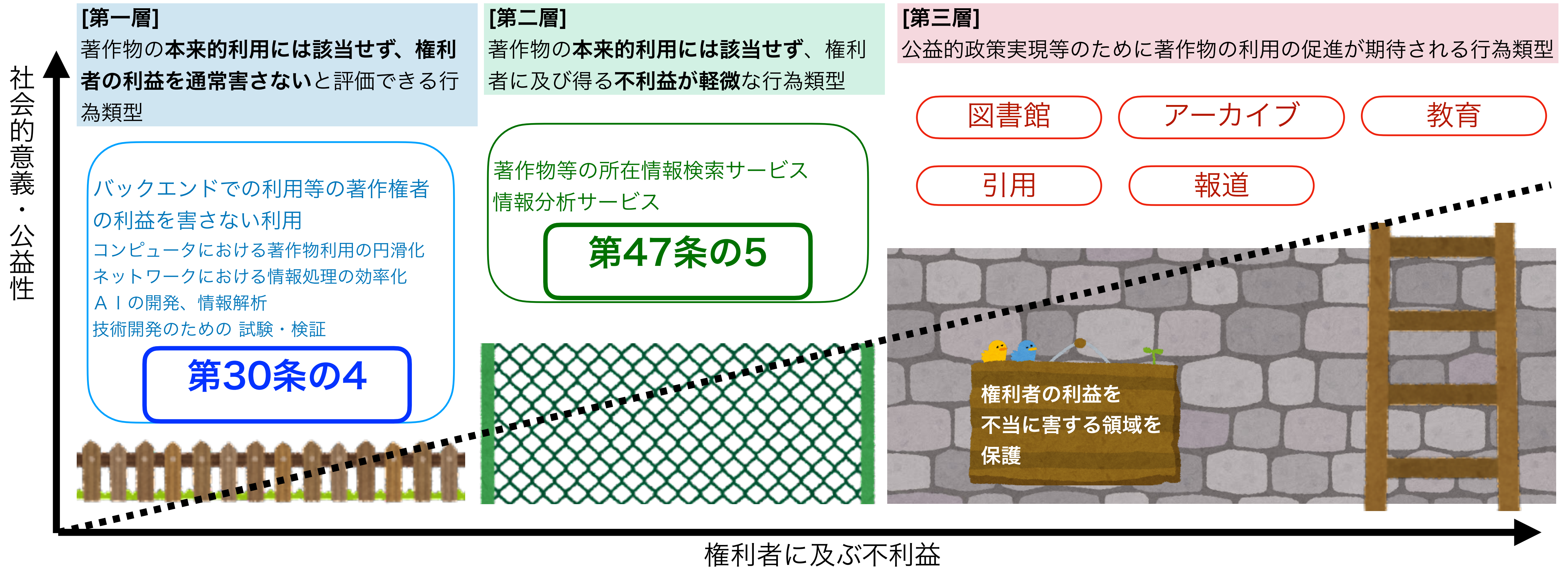
- 旧来は利用の目的や場面毎に具体的に権利制限を規定
- 自由な利用ができるケースを抽象的に定めた規定を階層的に追加
 - 柔軟性を高めると新たな利用行為に対応できるが、法規範が不明確になる可能性
 - 明確性と柔軟性の適切なバランスを備えた複数の規定の組合せによる多層的な対応

ベルヌ条約 スリーステップテスト (第9条より)

Step 1: 特別の場合であって

Step 2: 著作物の通常の利用を妨げず

Step 3: 著作者の利益を不当に害しない



* 文化庁「著作権制度の概要」資料より抜粋の上、改変

著作権法第30条の4

- 2018年改正著作権法第30条の4
 - 統計的解析を外すことで代数や幾何学的解析も明示的に
 - 「用に供する」への変更で情報解析用のデータセット販売も明示
 - 「記録又は翻案」から単に**利用**を認める
 - 思想又は感情の享受を目的としないならば情報解析を認める
 - 我々のような一般人からは下記が歯止めに見える
 - 必要と認められる限度において
 - 著作権者の利益を不当に害しない

2009(平成21)年改正著作権法 第四十七条の七

著作物は、電子計算機による情報解析（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、映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**を行うことをいう。以下この条において同じ。）を行うことを**目的とする場合**には、必要と認められる限度において、**記録媒体への記録又は翻案**（これにより創作した二次的著作物の記録を含む。）を行うことができる。(略)

2018(平成30)年改正著作権法 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 (略)

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、映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**

著作権法第30条の4 (続き)

- AIモデル学習時における第30条の4の解釈
 - ディープラーニング等の手法も問題ないことを明示
 - データセット購入、共有も可能と明示
 - 学習済みモデルの記録又は翻案以外の利用も可
 - 原則として日本では**著作権者の許諾なくAI学習が可能**
 - 非享受利用は著作権が及ぶところではないとの立て付け
 - 著作隣接権にも準用
- 第30条の4の制限
 - 「必要と認められる限度」
 - 傾向的にはデータが多ければ良い結果になるので必要と基本的に認められるだろう
 - 「思想又は感情を… 他人に享受させる」
 - AI学習は、著作物から特徴、特性を定量的な数値として抽出、蓄積していく作業であり、人間が著作物を視聴して知的・精神的欲求を満たす（表現を享受する）ことには向けられていない
 - 出力の目的(表現の享受)によっては**追加学習は該当しない場合**がある (が、学習段階での判定は難しい)
 - 「著作権者の**利益**を不当に害することとなる場合」
 - 著作物の利用市場もしくは潜在的市場と衝突するか* という観点から判断、米国フェアユースに似ている
 - 特に特定作家の画風(アイディア)を模倣という用途で議論がある (が、著作権の原則を考えれば難しい)

2018(平成30)年改正著作権法 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。

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、映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**

(参考) 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

https://www.bunka.go.jp/seisaku/chosakuken/hokaisei/h30_hokaisei/pdf/r1406693_17.pdf

2. 大量著作物による学習の妥当性

著作権法第47条の5

- ・ 情報処理及び結果の提供に付随する軽微利用
 - ・ 旧法のWeb検索サービスにおけるサイトのスニペットやサムネイルの提供の明示を拡大
 - ・ 「**軽微利用**」に該当すれば公表済み著作物を収集して検索的なサービスへ利用することが可能と明示
 - ・ 「利用に供される部分の**占める割合**、その利用に供される**部分の量**、その利用に供される際の**表示の精度**その他の要素に照らし軽微なもの」
 - ・ 思想享受目的(表現を享受)に該当するが、その利用が軽微であれば、AI学習(情報解析)において適用される

政省令で定める基準 一部抜粋要約 *

- ・ サービスに使用のデータベースに係る**情報漏洩防止の措置**を講ずること
- ・ サービスが改正法の要件に適合したものとなるよう、事前に**学識経験者に対する相談**等の必要な取組を行うこと
- ・ インターネット情報検索サービスを行う場合、**ID・パスワードにより受信が制限された情報**や、**業界慣行に沿って情報収集禁止措置**がとられた情報を使用しないこと

2018(平成30)年改正著作権法 第四十七条の五

電子計算機を用いた情報処理により新たな知見又は情報を創出することによって著作物の利用の促進に資する次の各号に掲げる行為を行う者（当該行為の一部を行う者を含み、当該行為を**政令で定める基準**に従って行う者に限る。）は、公衆への提供等（略）が行われた著作物（略）について、当該各号に掲げる行為の**目的上必要と認められる限度**において、当該行為に付随して、いずれの方法によるかを問わず、利用（略 以下この条において「**軽微利用**」という。）を行うことができる。ただし、当該公衆提供等著作物に係る公衆への提供等が著作権を侵害するものであること（略）を知りながら当該軽微利用を行う場合その他当該公衆提供等**著作物の種類及び用途並びに当該軽微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 電子計算機を用いて、検索により求める情報（略）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（略）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。

二 **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。

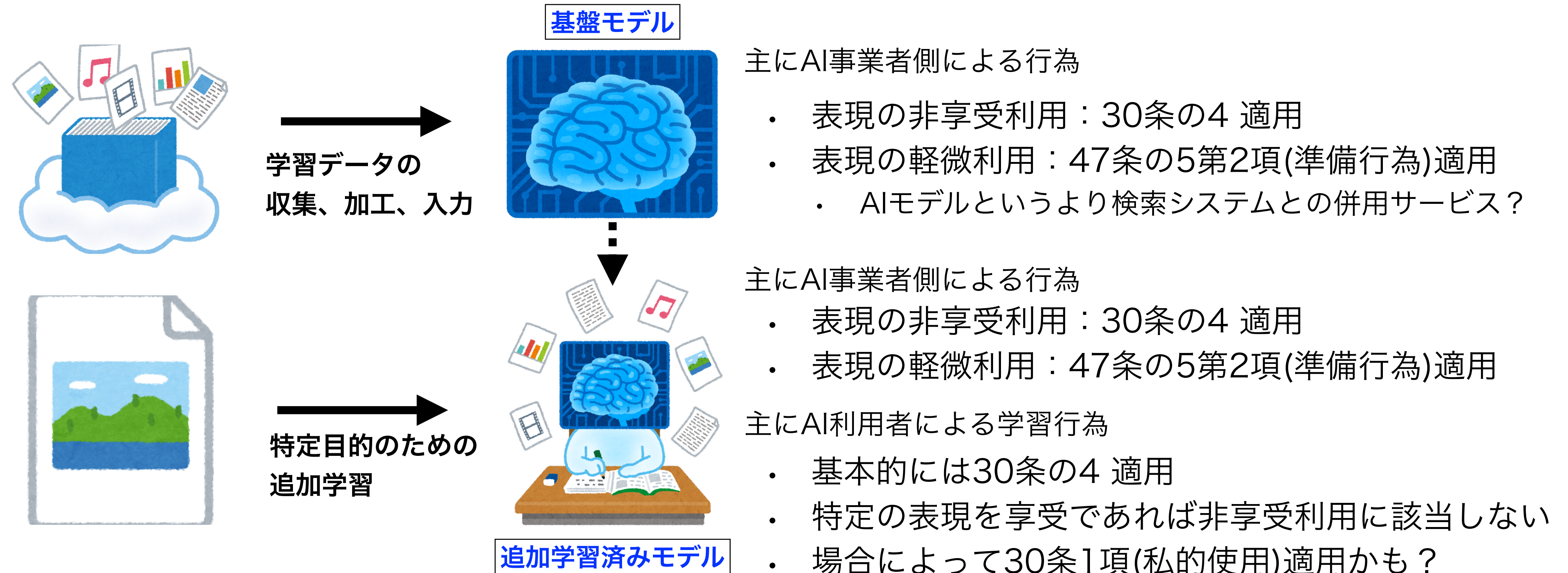
三 略

2（略）は、公衆提供等著作物について、同項の規定による**軽微利用の準備のため**に必要と認められる限度において、**複製若しくは公衆送信**（略）を行い、又はその**複製物による頒布を行う**ことができる。ただし、当該公衆提供等**著作物の種類及び用途並びに当該複製又は頒布の部数及び当該複製、公衆送信又は頒布の態様に照らし**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。

(参考) 平成30年著作権法改正に伴う政省令改正の概要

https://www.bunka.go.jp/seisaku/chosakuken/hokaisei/h30_hokaisei/pdf/r1406693_05.pdf

AI開発・学習時の法定利用行為



- 「表現」を享受させる目的がなければ30条の4、軽微利用なら47条の5
 - AI事業者による一般的なAIモデル開発は基本的に30条の4適用がほとんどだろう
 - 著作物の創作的「表現」が結果物へ影響することを目的としているなら適用外の可能性
 - 特定作品の表現が強く影響することを目的とした限られた数量の著作物による学習

合衆国法典第17編第107条

- 米国フェアユースは情報解析利用に留まらない包括的権利制限規定
 - 要件：
 - (1)使用の目的および性質(商業性がある**営利目的**かを含む)
 - (2)著作権のある著作物の性質
 - (3)著作権のある著作物全体との関連における使用された部分の量および実質性
 - (4)著作権のある著作物の**潜在的市場**または価値に対する使用の影響
 - (1)、(4)の要素での影響が大きいとされている
 - Google v. OracleのJava API訴訟では(2)、(3)も含めた4要件全てを満たすと判断
 - Google Books判決 (全米作家協会 v. Google)
 - 変容的利用の理論：利用された著作物の本来の目的から見て、その利用がどの程度変容的かで判断
 - 全文検索提供のための書籍複製やスニペット表示について**変容的利用の理論**に基づきフェアユースを肯定
- 基盤AIモデル開発における学習での著作物利用
 - 変容的利用に該当する可能性が高いと**一般的にフェアユースと判断**されているのだろう
 - そもそも判例法の国であり、判例が出るまで突っ走る
 - ただ、**市場の変化**によって一般的な判断が変わる可能性があり、最終的な判断は訴訟に委ねられる

EUデジタル単一市場著作権指令

- 第3条 学術研究目的でのテキストおよびデータマイニング
 - 学術研究機関が第三者著作物を使用のTDMを行う際、著作権者の排他的権利が及ばないと規定
 - AIモデル開発にも適用されると解されている
 - 契約等で3条適用のTDMを制限することはできない (第7条、前文18)
- 第4条 テキストおよびデータマイニングのための例外または制限
 - 学術目的以外を含め、第三者著作物の複製および抽出はTDMの目的に必要な期間、保持できる
 - オンラインで提供のコンテンツは機械可読な手段での意思表示の場合のみ権利留保 (前文18)
- 基盤AIモデル開発における学習での著作物利用
 - 4条適用での基盤モデル開発はオプトアウト処理が必要
 - 学術研究目的で3条適用のAI学習の成果を営利企業が商用利用することは可能
 - AIモデル開発主体を産学協働体制にすることで一部企業の独占を排除する狙い？
 - モデルのライセンスを学術研究目的に制限しているのもその流れ？

権利制限の各国比較

- 日本はAIモデル開発がやりやすいパラダイスと称する意見もある
 - 著作権法観点では包括的かつ明示的な権利制限が存在し、世界で最もAI学習に対しての障壁が低い
 - ただし、実質的には欧米と大差はなさそう

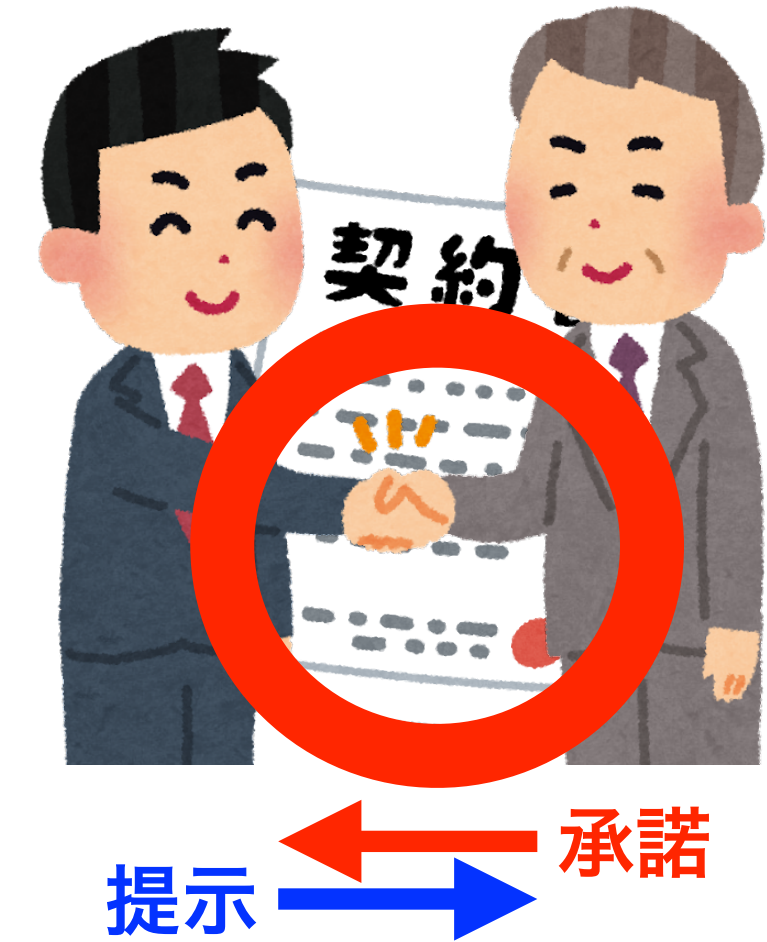
	営利利用	著作者の許諾	備考
 日本	可能	不要	著作者の利益を不当に害しない限りは制限がない
 米国	可能	不要	市場の変化によりフェアユースと認められない可能性
 EU	営利目的に制限	営利の場合は権利者が利用を明示的に制限可能	学術研究目的にあれば権利者の許諾を必要としない

• 準拠法問題

- 著作権法に関してどこの国の法律が適用されるかは「**当該著作物が利用される地**」(利用行為地)
 - 諸外国の者の著作物を利用した日本国内でのAIモデル開発には基本的に日本法が適用
 - 利用行為地の判定：
 - 開発者がいる国、企業所在地、サーバーの所在地等が異なる場合は不明瞭になるケースがある
 - 米国、欧州企業が日本でAIモデル開発を行うケースが出てくる可能性

著作者側からAI学習を止めさせる

- 著作物の権利者が日本国内でのAI学習を止めさせるには？
 - 日本国内に適法に持ち込まれたデータによる学習は問題ない
- 「自分の作品のAI学習への利用禁止」と表明があった場合：
 - 著作権法の権利制限を覆すオーバーライド条項
 - 単なる**表明では無効**
 - ライセンスによる許諾方式では禁止できない可能性が高い
 - 少なくとも権利者と利用者間の**契約成立が必要**
 - 契約に合意していない者は縛られない
 - 定型約款方式でも問題ない
 - 利用規約の場合、**明確な提示と承諾**のフローが重要
 - 本来的に非享受目的の利用には著作権が及ばない
- 欧州では学術研究目的でのTDM禁止は無効
- 米国では禁止の条項は有効と考えられる
 - Bowers v. Baystate：リバースエンジニアリング禁止の契約を有効と判断

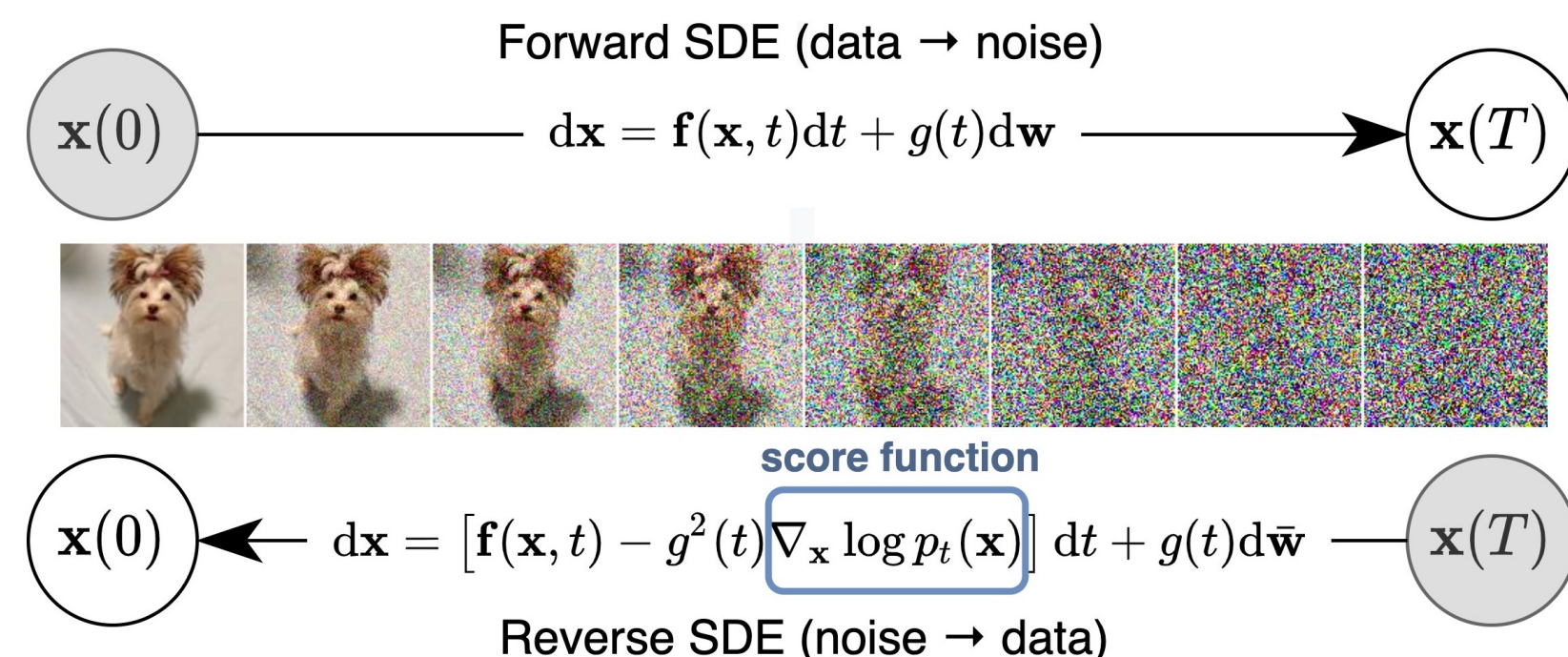


(参考) 海外におけるデザイン・ブランド保護等新たな知財制度上の課題に関する実態調査

https://www.meti.go.jp/policy/economy/chizai/chiteki/pdf/reiwa3_itaku_designbrand.pdf

各地で起きる訴訟

- GitHub Copilotに対する集団訴訟
 - 原告側は著作権を基にしたオープンソースライセンスへの違反という訴え
 - 帰属表示(作者名)や著作権表示、ライセンス条項を表示する条件へ違反
 - 被告側はAI学習はフェアユースに該当と主張
 - フェアユースは著作権への権利制限であり、ライセンス効力は及ばない
- Stability AI、Midjourney、DeviantArtに対する著作権侵害集団訴訟
- Getty ImagesによるStability AIに対する著作権侵害訴訟
- 全ての訴訟が結局は無許諾利用の訴えをフェアユースで抗弁の構図になるのでは？
 - Google Books判決の前例に倣えばフェアユースを否定するのは難しく見える
 - 拡散モデルを含む主要基盤モデルが学習の解析結果として**アイデアを保持しているのか？それとも表現と呼べるものを保持しているのか？**という問いへの回答が出るかも？



閑話：AI規制、法制

- ・ 著作権以外の問題も多い
 - ・ プライバシー保護、政治的中立性、透明性、虚偽情報、その他の社会倫理
 - ・ 世界各地でAI規制の機運

・ ハードロー路線のEU

- ・ **EU AI規則が発効間近 (6/14 欧州議会可決)**
 - ・ AI出力がEU域内で利用される場合に適用され、違反にはあのGDPRよりも高額な罰金
 - ・ 高リスクAIは、データ品質、透明性確保、生成ログ保存義務、人間による監視、堅牢性と正確性、GDPR等への準拠等の厳しい規制と義務
 - ・ 基盤モデル、生成AIにも高リスクAIに準じる規制。特に**著作物利用の透明性確保**

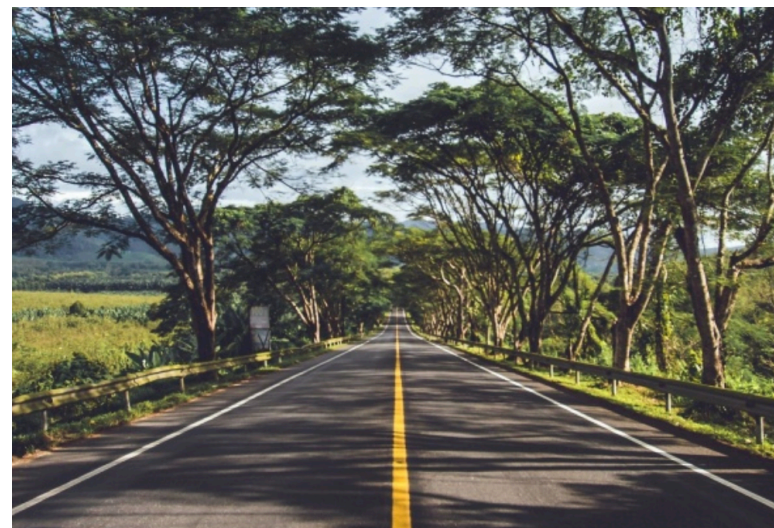
・ ソフトロー路線の米国

- ・ AI権利章典とAIリスク管理フレームワークの原則に沿って、各官庁のガイドラインで対応
- ・ ホワイトハウスのリーダーシップの下で専門家の検証と関係企業の統制

・ ノーロー路線の日本

- ・ EUと米国は手法は異なるものの基本的人権と民主主義の保護で似た価値観を共有
- ・ 日本はAI開発に支障がないことを優先し、制約をなるべく作らない方向？


舗装路




林道




獣道



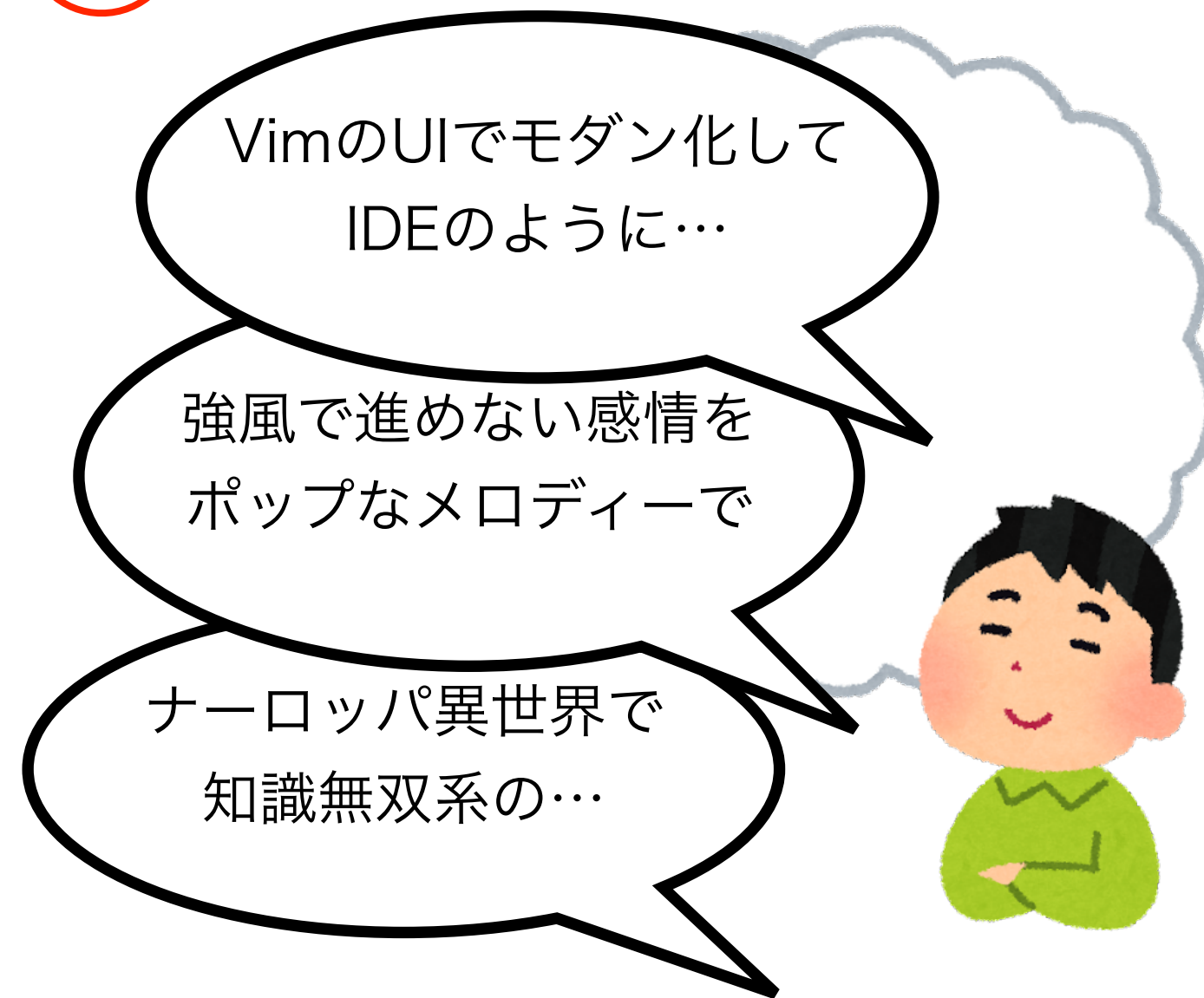
3. AI生成物の著作物性

著作物の定義 (再掲)

- 思想又は感情を創作的に表現したものであって、文芸、学術、美術又は音楽の範囲に属するもの (日本国 著作権法第2条第1項第1号)

思想又は感情

人による何らかの考え、気持ち



創作的

人の何らかの個性を反映



表現したもの

人の五感で感知し得る程度に具体的な表現



文芸、学術、美術又は音楽の範囲に属する

言語、音楽、舞踊・無言劇、美術、建築、図形、映画、写真、プログラム、編集著作物、データベース著作物

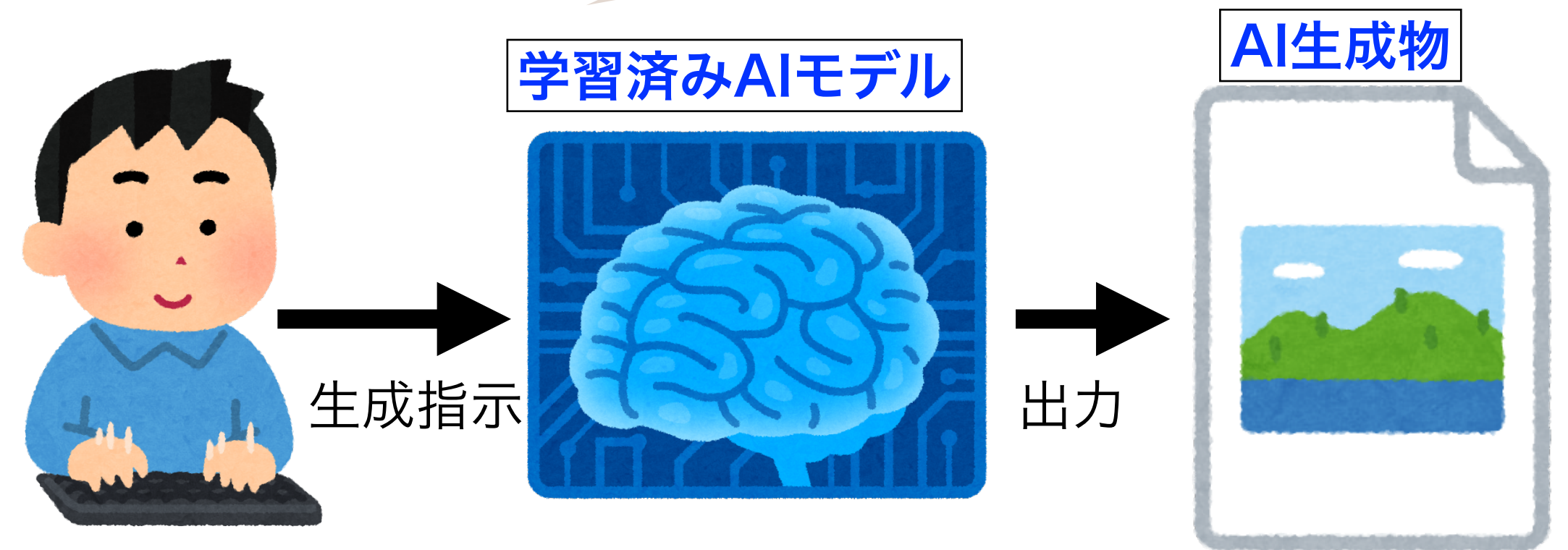
「人の文化的活動で生み出された物」にあたればよい

AI生成物は人の創作の産物なのか？

- シンプルな指示による生成AIの出力 🇯🇵

- 「思想又は感情」という精神活動を最終作品で表現されているか？
- 人の**創作的意図と創作的寄与**はあるのだろうか？
 - 人ではないAIモデル自身が表現へ寄与している
 - よって、**生成AIの自律的出力に著作物性はない**
 - 創作的寄与のハードルを低くする議論もあるが、
 - ただし、まだはっきりとはしていない

人間への委託だとどうなるか？



- 日本以外の法域

- ほとんどの法域の著作権法の下で、**著作権の発生は人間の創作物に限られる**

- 🇺🇸 米著作権局：「AI技術が出力の表現要素を決定する場合、生成された素材は人間の著作物ではない」
 - 「人間が作品の**表現に対して創造的な制御**を行使し、伝統的な著作権要素を「**実際に形成**」していたか」

- 例外：英国法の影響下にあるアイルランド、香港、ニュージーランド等

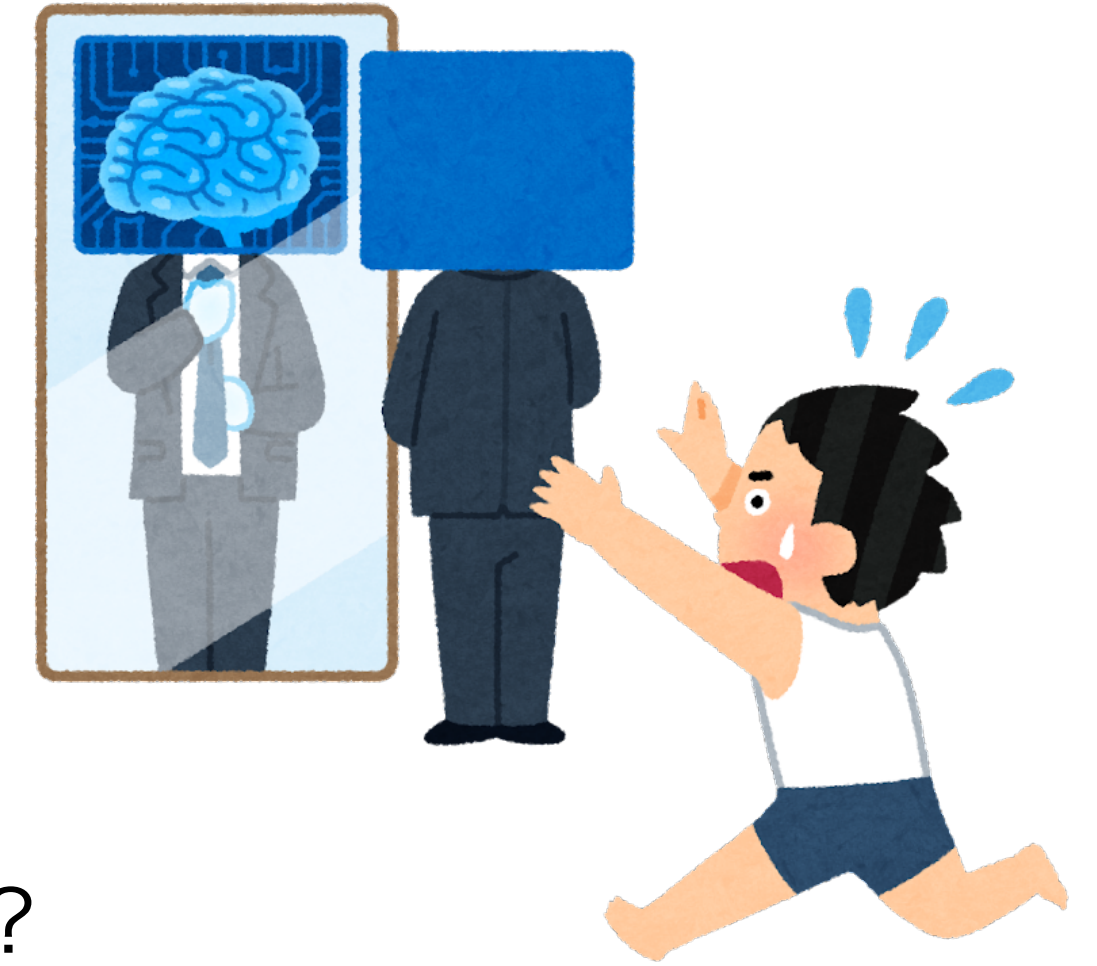
- 🇬🇧 英国1988年著作権、意匠および特許法(CDPA) 9条(3)

- 「コンピュータにより生成される文芸、演劇、音楽又は美術の著作物の場合には、著作者は、**著作物の創作に必要な手筈を引き受ける者**であるとみなされる」

- 人の関与が全くない状況で適用されるかは不明瞭であり、かつ各国から批判が多い
- AI使用者、AI所有者、AI開発者、AIそのものの内、誰が著作者となるかが不明瞭

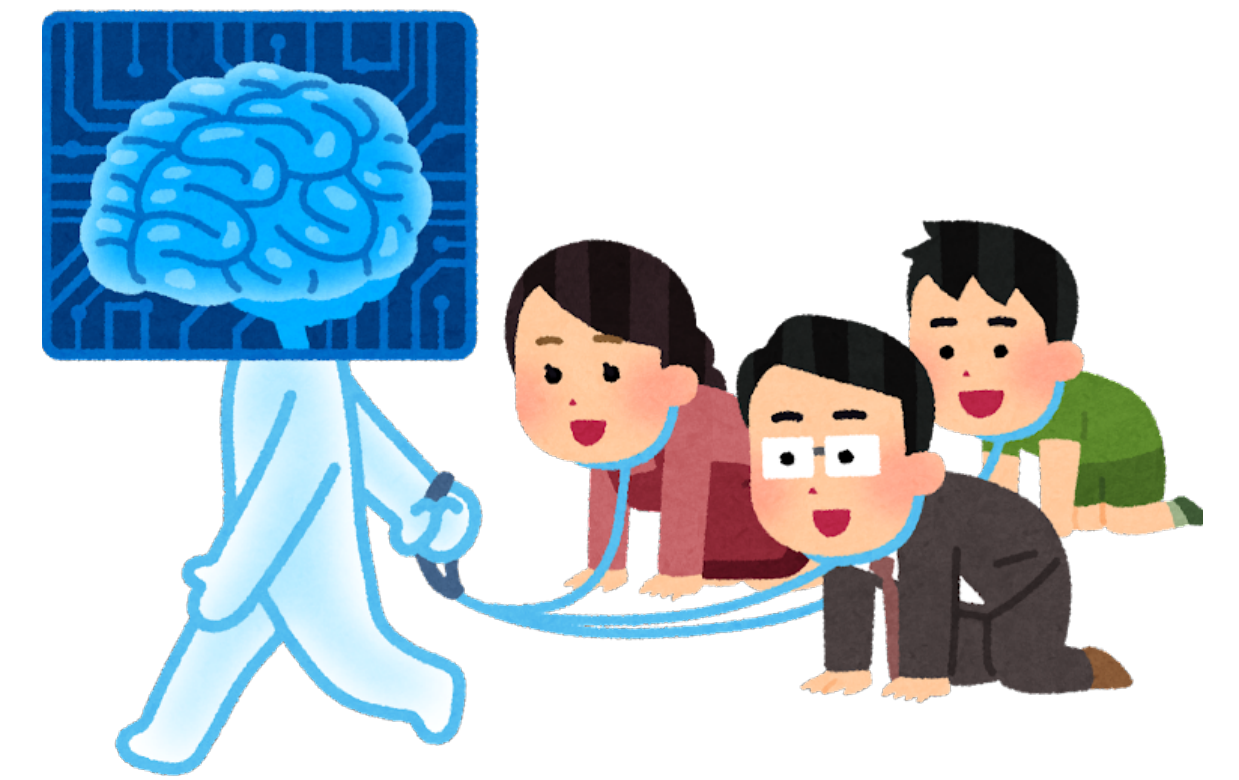
著作物性否定による問題

- 生成AIによる自律的な出力には指示者の著作権が及ばない
 - 著作権的には自由な複製、翻案、改変が可能と考えられる
 - 膨大な量のパブリックドメインコンテンツが発生
 - 権利帰属の不明瞭さによる混乱、クリエイター市場の停滞が発生するかも
 - 第三者もしくは自身によるAI生成物を人間の創作とする詐称が横行？
 - AI生成物が広く流通することで人間による二次的著作物の権利が狭まる可能性？
 - 人間の創作の**制作過程**に対して**透明性、記録性**が求められる可能性？



AIに仕事を奪われる人間

- かといって、AI生成物に著作物性を認めると…
 - 英国CDPA 9(3)問題と同様に**著作者は誰？問題**が発生
 - AIユーザー以外にAI開発者、AI所有者、AI自身に隣接権を含む権利の可能性
 - 創作への寄与という観点ではAI開発者に権利が発生してしまうのでは？
 - AIによる著作物の氾濫が発生
 - AIという**計算機パワーを持つ者がコンテンツ市場を独占**する方向へ？



AIの奴隷にされる人間

いずれにせよ著作権観点ではカオスな時代が到来する

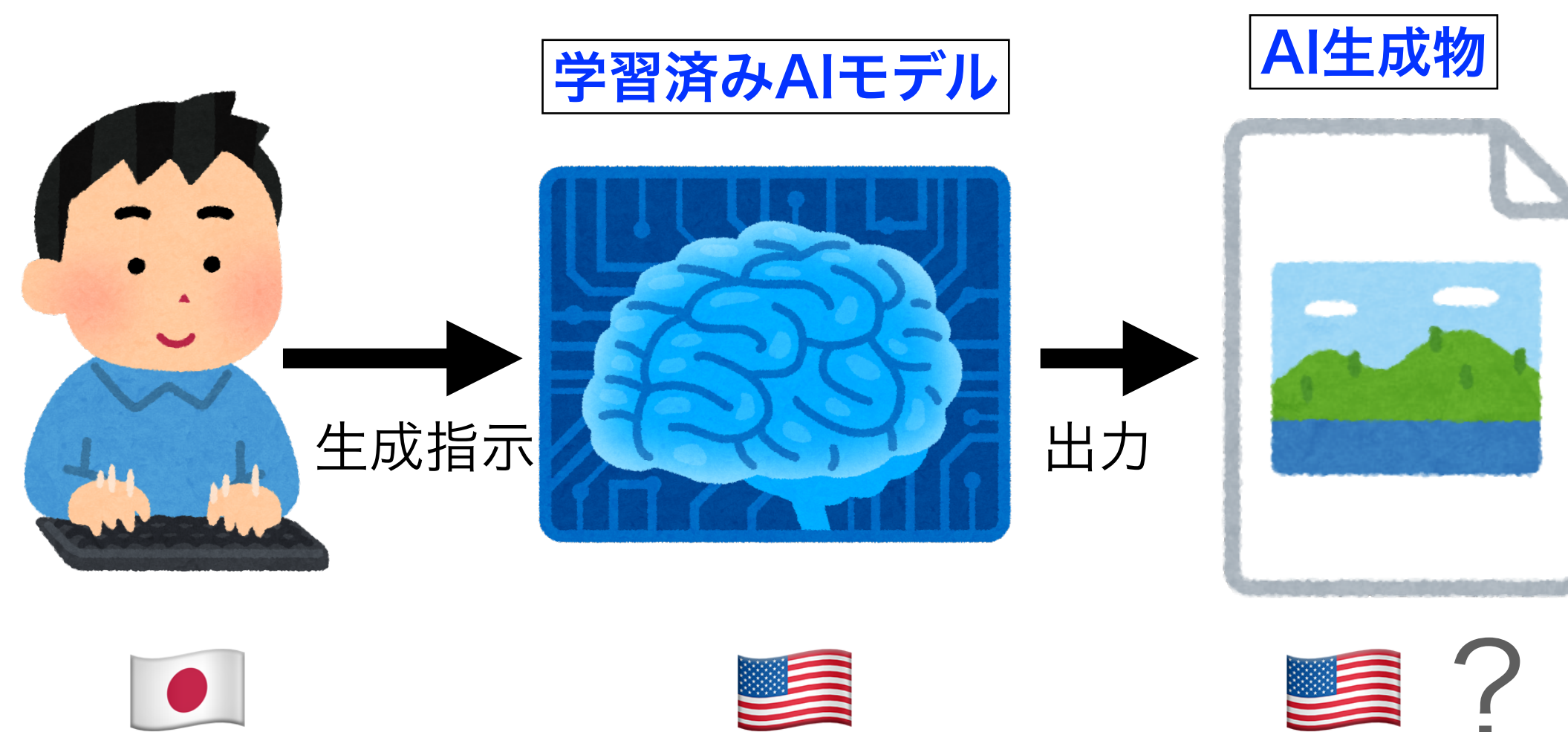
国境を越えるAI利用

- ・ 準拠法問題

- ・ 著作権法に関してどこの国の法律が適用されるかは「当該著作物が利用される地」(利用行為地)
 - ・ AI生成物を創作し、出力するのはAIモデル
 - ・ AIモデルが動作する地、運営されている地が行為地と考えられる

- ・ AI利用者の所在地とAIモデル動作地が異なる場合にどうなるか？

- ・ AIモデル動作地で著作権が発生
- ・ AI生成物に対して著作権発生要件が異なることで問題が起きる可能性
 - ・ 積極的に著作権を認める法域から認めない法域のサービスを利用すると著作権が発生しない？
 - ・ 保護国法説を取るか、保護期間の相互主義を取るか



AIで著作物を出力できないの？

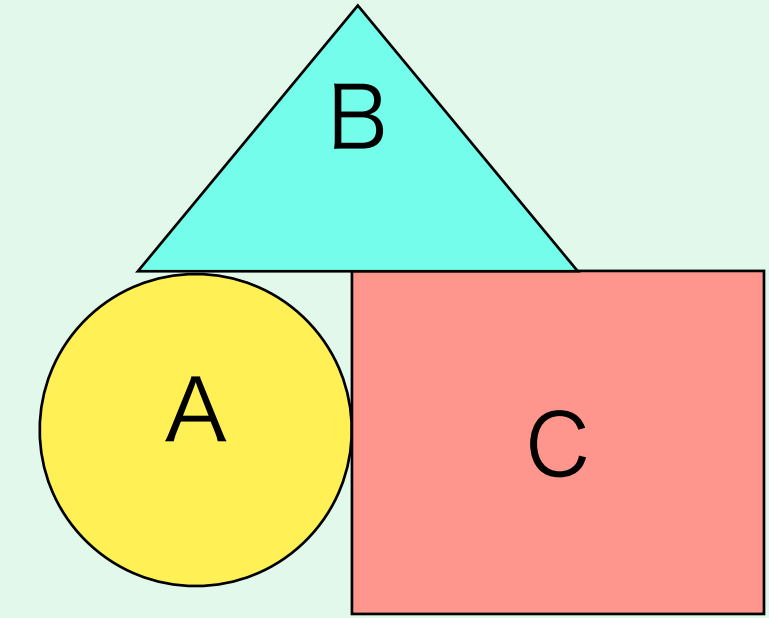
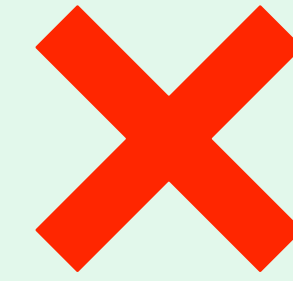
- どうにかして生成AIにて確実に著作物を出力させられないか？
 - (1) 非常に長く、具体的で詳細なプロンプトを入力した場合は？
 - アイディアをAIへ提供しているだけに過ぎず、**指示者の精神性が表現されていない**と判断される可能性
 - ソフトウェア開発請負では設計・仕様書を書いた者ではなく実装者に原始的著作権
 - プロンプトに著作権が発生する可能性はあるが、**プロンプトと出力は直接関係しない**
 - プロンプトが著作物と仮定しても、出力の創作性に関与するのはAIであり、二次的著作物には該当しない
 - (2) 何度も繰り返して目的のコンテンツが得られるまで同じプロンプトを入力する場合は？
 - 個々の出力の著作物性は(1)と同じ
 - 目的のコンテンツを選択する行為を創作的寄与と看做す論がある
 - 特に日本に見られる考え方であるが、主流の論にはならないと考えられる
 - (3) 目的のコンテンツが得られるまで少しずつプロンプトを修正し続ける場合は？
 - 基本的に(2)と同じ
 - 修正の繰り返しが指示者の「思想又は感情」の表現と看做せる余地が若干あるかも？
 - ただし、これも日本で見られる傾向

生成AIの出力を著作物として扱うハードルは高い

AI生成物を組み合わせてみては？

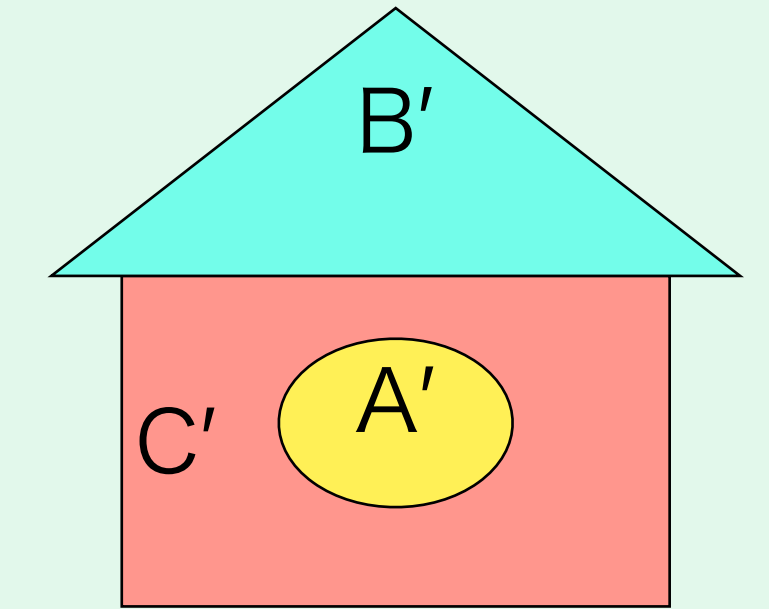
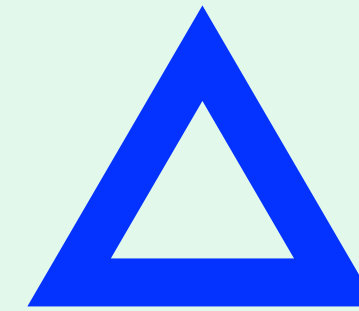
・ パターン 1

- ・ AI生成物A、B、Cを単純に組み合わせるとどうか？
 - ・ 単にA、B、Cを組み合わせただけでは著作物とはならない
 - ・ 組み合わせに人の「思想又は感情」を見出すのは難しい



・ パターン 2

- ・ AI生成物A'、B'、C'の組み合わせを工夫するとどうか？
 - ・ 人の手で組み合わせ、思想が表現されていれば著作物性があるかも
 - ・ 選択と組み合わせが限られていればやはり難しい

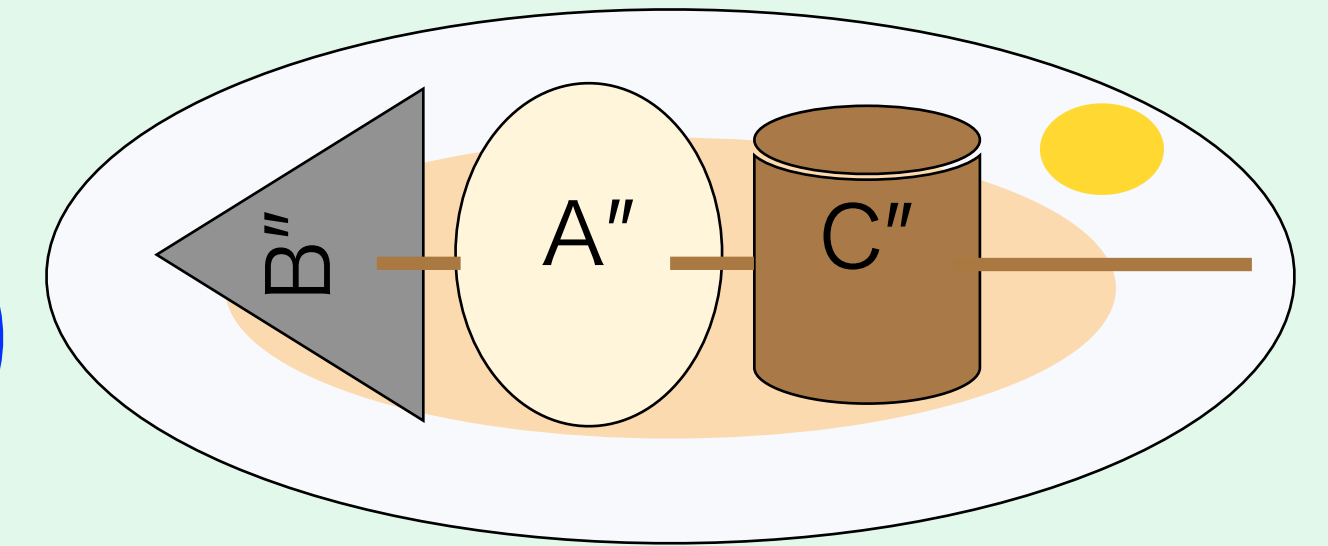
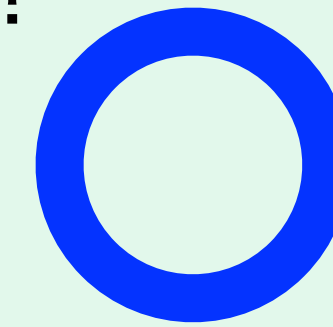


- ・ 単なる組み合わせ作業では創作性があると見做すことは難しい

組み合わせ以外の工夫はどうか？

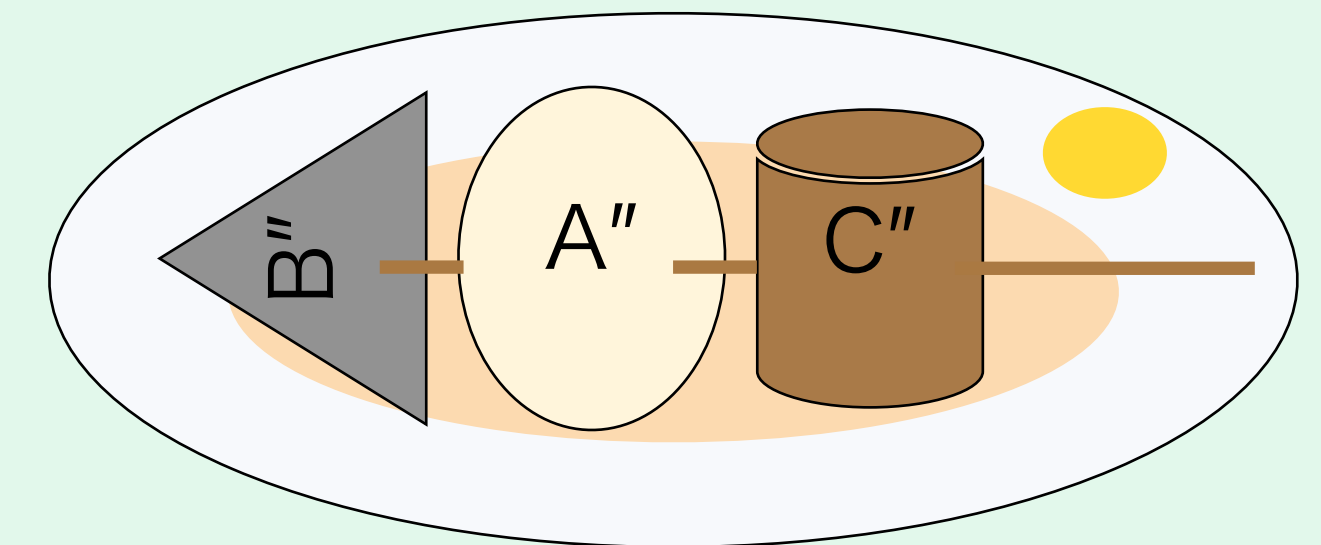
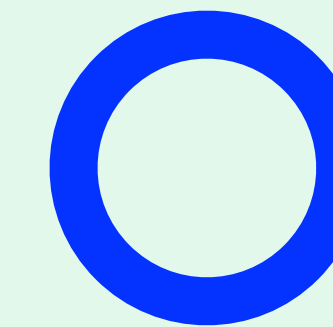
・ パターン 3

- ・ AI生成物A''、B''、C''を十分な人の手で繋ぎ合わせるとどうか？
 - ・ 人の手で**接合する部分に十分な創作性**があれば全体に著作物性
 - ・ 実装のための**表現手法が複数存在し、表現が凡庸でなければ良い**
 - ・ ただし、A''、B''、C''に著作権は発生しない



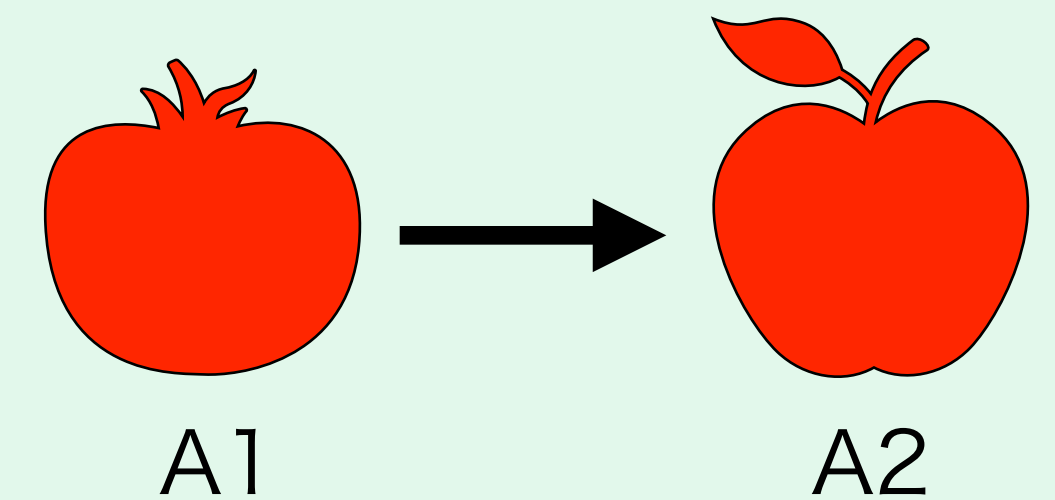
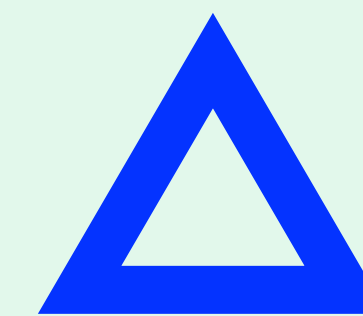
・ パターン 4

- ・ 人間の作品の一部でAI生成物を利用した場合はどうか？
 - ・ パターン 3と考え方は変わらない
 - ・ **人の手による作品部分の創作性が重要**



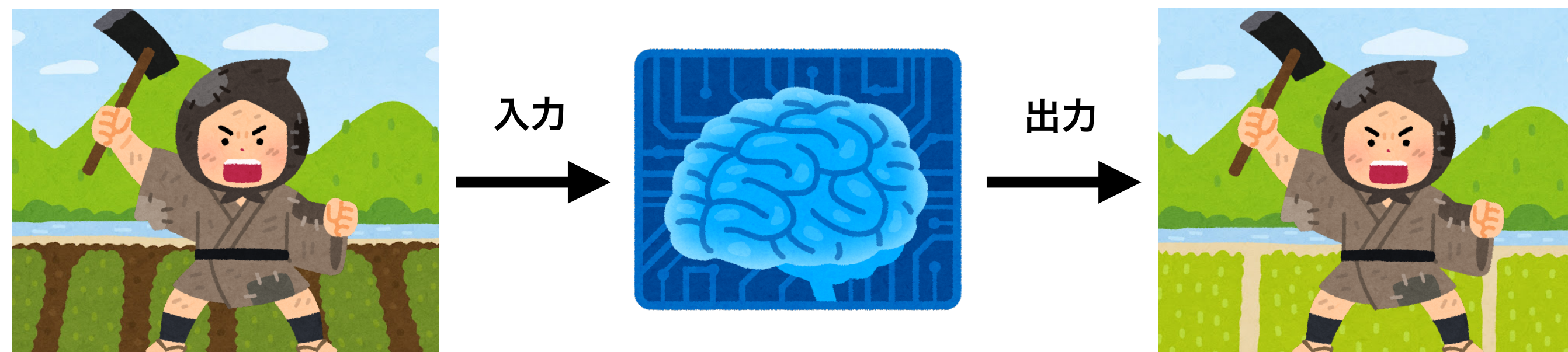
・ パターン 5

- ・ AI生成物A1を改変した場合はどうか？
 - ・ **人間による創作性のある十分な改変、修正**が加わればA2は著作物
 - ・ 表現の根幹となる処理に**人間の個性が反映**される必要
 - ・ 改変部分だけが著作物性があるものとして扱われる



AI生成物の著作物性

- ・ シンプルな指示入力の場合
 - ・ 現状では**人の手による創造的作業がある場合に著作権が発生すると考えたほうが良い**
 - ・ 軽微な修正、調整程度では人の創作性は認められない
 - ・ 確実に著作権の発生がない部分を含むコンテンツが増大、著作権管理が必要になるかも
 - ・ 人の手による十分に創造的な作業があれば、少なくとも誰が著作権者であるかという問題は防止される
 - ・ AIを利用した創作の権利主張には**創作過程の記録と透明性の確保が重要**となる可能性
- ・ 修正、改変、改作等を意図した著作物を伴う入力だとどうか？
 - ・ 改変された部分に人間の創作的寄与があるとは認められないだろう
 - ・ ただし、入力著作物の表現(本質的特徴)が残るなら入力の「複製*」に該当
 - ・ 入力された著作物に対する**原著作者の権利が出力生成物に及ぶ**
 - ・ つまり、AIを通せば著作権が消え去るわけではなく、利用には許諾が必要



(注) 著作物の複製とは、既存の著作物に依拠し、その内容及び形式を覚知させるに足りるものを再製することをいうと解すべきである
最判 昭和53年9月7日

AI生成物の僭称問題

- 今後、生成AIの受容で著作物性の僭称が大きな問題となる可能性
 - AI生成物を人間の著作物と僭称する動機は幾つもある
 - 人間の創作であることを誰がどのように証明するのか？
 - 原則では著作権者が著作物性を証明する必要があるが、僭称しない者まで巻き込む
 - 「スレイヤーズ」の挿絵担当イラストレーターへのAI疑惑騒動は僭称問題の幕開けを示すかも
 - 僭称の禁止化はどうか？
 - 著作権法121条：著作物の僭称を禁止 (著作物の複製物を頒布する場合)
 - 非著作物を著作物として僭称することを禁止？
 - 現在の市場への影響が大き過ぎ、無方式主義の著作権制度と矛盾するかも？
 - AI生成物に著作権とは異なる知的財産保護はどうか？
 - 英国1988年CDPA 9条(3)、EUデータベース権あたりが参考になるかも
 - 生成のための手筈を整えた者へ著作権よりも限定された独占権という考え方は可能かも？
 - CDPA 9条(3)の開発者、利用者、所有者のうち誰が権利を持つべき論争と同じことが発生しそう
 - 結局、計算機パワーを持つ者が独占する世界へ向かうのでは？

GitHub Copilotの出力の著作物性



- GitHub Copilot (準拠法：米国カリフォルニア州法 🇺🇸)

- OpenAI Codex、GPTを利用した開発支援、コード提案ツール

GitHub Copilot Product Specific Terms (2023.3.29)

2. Ownership of Suggestions and Your Code.

GitHub does not claim any ownership rights in Suggestions. You retain ownership of Your Code.

(GitHubは、出力されるコード提案の所有権を主張するものではありません。お客様のコードの所有権は、お客様に帰属します。)

- 途中までのルーチンの残処理部分を自動補完
 - 外形的にはIDEの自動補完の延長に見える
 - IDE自動補完は定型的な形式の入力を支援するが、Copilotは本質的なコード表現まで含む
 - ただし、人間のそれまでの入力表現に影響を受け、人に主導権のある選択と修正を必要とする
 - 生成直前のコードの「複製」として著作権が及ぶ可能性？
- コードのコメントから実行可能コードを生成
 - 生成コードに著作権は発生しない
 - ただし、
 - 生成コードをそのまま利用せず、既存コードに適合するように修正すれば著作物性がある
 - それなりの規模の開発であれば、生成コード由来の処理は一部に過ぎず、複雑な組み合わせとなる
- **現時点**において著作物性の有無を然程は気にする必要はない
 - ただし、個々のCopilotの提案部分は、利用者が所有する著作物ではないだろう

GitHub Copilotの出力の著作物性 (続き)



- ただし、ソースコードを外部へ提供する場合は注意が必要かも？
- 著作権のないコードを外部へ自身の著作物として提供した場合、
 - 虚偽の著作権宣言は違法であるが、権利者がいなければ誰の権利も侵害しない
 - 僭称問題に関しては倫理的な問題が大きいかも？
 - 実際の可能性としては無視できると考えられる
- オープンソースプロジェクトへ貢献する場合、
 - Contributor License Agreement(CLA)とDeveloper Certificate of Origin(DCO)
 - 貢献者がコードに対して**完全に知的財産権の権利行使できる立場**であることを表明し、著作権と特許のライセンスをプロジェクト側へ付与する
 - 著作権のないコードの存在によってエコシステムに混乱が生じる可能性
 - ただし、現時点で問題となっている事例はない

ChatGPTの出力の著作物性



- ChatGPT (準拠法：米国カリフォルニア州法 🇺🇸)
 - GPTを基盤に開発されたAIチャットボット
 - 人間同士の対話を模倣し、極めて汎用的かつ即興的な対話を実現する
 - 小説や詩を創作し、ゲームをプレイし、プログラミングも行う
- 出力の著作物性
 - プロンプトを入力し、出力として創作物を得たとして
 - 「創作的」に「表現」するのはAIであるので、出力に著作権は発生しないと考えべきだろう
 - 凝ったプロンプトだろうが、出力を組み合わせようが、著作物性はないと考えられる
 - 著作権の主張にはパターン3, 4, 5のように十分な人の手による創作を加える必要
 - ただし、
 - 出力結果を誰も知らないのであれば誰にもAI生成だと分からない
 - 著作権のない作品に対しての著作権保有の宣言の扱いは不明瞭
 - 不正競争的な観点あるいは倫理的な問題として扱われるか？

OpenAI Terms of use (2023.3.14)

Subject to your compliance with these Terms, OpenAI hereby assigns to you all its right, title and interest in and to Output. This means you can use Content for any purpose, including commercial purposes such as sale or publication, if you comply with these Terms.

(お客様が本規約を遵守することを条件として、OpenAIは出力に関する全ての権利、権原および利益をお客様に譲渡します。つまり、お客様は、本規約を遵守すれば、販売や出版などの商業目的を含め、いかなる目的にもコンテンツを使用することができます。)

イラスト生成AIの出力の著作物性

• イラスト生成AI

- DALL-EやMidjourney等のクラウドサービス
- オープンに頒布されるStable Diffusion (準拠法：モデル実行地)
 - 追加学習された多くの派生モデル
 - プロンプトからの画像生成、入力した画像を基にした修正画像生成(img2img)



UFOにさらわれる雌牛
(Public Domain)

• 出力の著作物性

- ChatGPTと同様、プロンプト入力を基にした出力には著作物性はないと考えられる
 - 著作権の主張にはパターン3, 4, 5のように十分な人の手による創作を加える必要
 - ただし、画像の部分的修正等の著作物翻案用途であれば、Photoshopのような道具を使用したと同義かも？
 - 利用者が入力元画像の著作者であれば、本質的特徴が残ることを前提に創作過程の一部であると考えられる
 - テキスト作品やプログラムと異なり、AI生成物に対して二次的著作物へ昇華させるインセンティブが少ない
 - 既にイラストサイト、マーケットプレイスでAI生成イラストが氾濫している現実

CreativeML Open RAIL++-M License

6. The Output You Generate. Except as set forth herein, Licensor claims no rights in the Output You generate using the Model. You are accountable for the Output you generate and its subsequent uses.

(6. お客様が生成した出力。本規約に定める場合を除き、ライセンサーは、お客様がモデルを使用して生成した出力についていかなる権利も主張しません。お客様は、生成した出力およびその後の使用について説明責任を負うものとしします。)

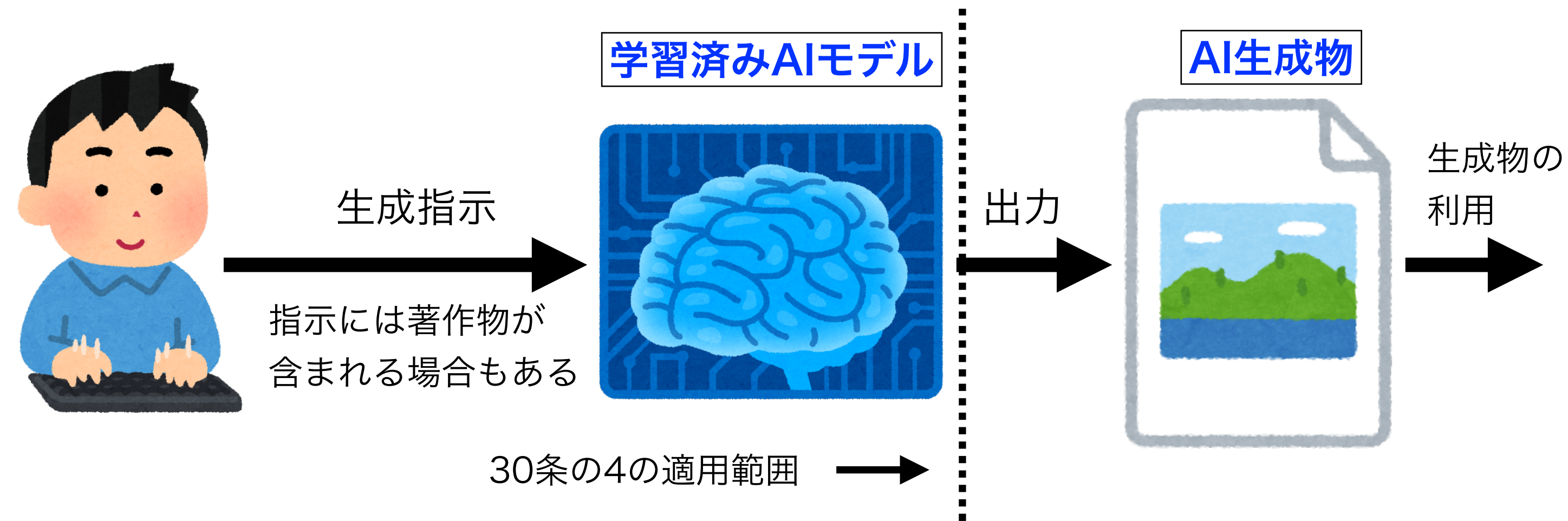
4. 他者著作物に対する著作権侵害

著作権侵害の基本的考え方



- 著作権は「思想又は感情を創作的に表現した」著作物を保護する
 - データ(事実)は含まれない
 - アイディア(作風、画風、キャラクター、アルゴリズム)も含まれない
 - 創作性のない(凡庸な表現、アイディアと表現が不可避で一致等)ものも含まれない
- 生成AIによる出力でも著作権侵害の成立には以下の4要件を満たす必要
 - 1. 著作物であること
 - **創作性**がなければ侵害とはならない
 - 2. 著作権が存在していること
 - 著作権が消滅しておらず、著作権法で保護される著作物である必要
 - 3. 著作権の効力が及ぶ範囲で利用されていること
 - 複製権、翻案権について侵害が成立するためには**依拠性と類似性**が必要
 - 4. 利用者が著作物利用について正当な権原を有していないこと
 - 著作権者から許諾を得ていない状態

利用・生成時の法定利用行為



- ・ 利用行為の状況によっては下記権利制限に該当する可能性もある
 - ・ 私的複製 (30条1項)
 - ・ 検討過程利用 (30条の3)
 - ・ 教育機関利用 (35条)

・ 生成指示の入力時

- ・ 指示に著作物が含まれた場合は基本的に30条の4適用と考えられる
 - ・ 学習時と同様に「情報解析」行為として適法
 - ・ 例外は思想享受目的、特に指示+第三者著作物の入力
 - ・ 現実としては指示行為が即座の権利侵害に結びつくことはあまりなさそう

・ AI生成物の出力およびその利用時

- ・ 生成物出力以降の工程には30条の4は適用されない
- ・ 生成物に**既存著作物**の創作的表現が含まれるのであれば権利侵害の可能性が生じる
 - ・ AIサービス提供者による出力であれば軽微利用 (47条の5第1項2号) 該当の可能性が残る

生成AI特有の依拠性判断

- AI生成物による他者著作物への権利侵害は**通常**の著作物と同様の判断
 - 「創作的」な表現が共通する必要がある、創作性がなければ否定
 - 「類似性」の判断については、通常^の著作物における判断と変わらない
 - ただし、「依拠性」の判断は、利用者自身による依拠に追加して**生成AI特有の依拠性判断**
- 既存著作物が学習用データセットに含まれていた場合の依拠性の考え方
 - **否定説**：学習済みパラメータには既存著作物が全く残っていないので認めるべきでない
 - **肯定説**：モデルの学習に利用されたのであれば認めるべき
 - 既存著作物に依拠したかを厳密に特定しない判例や既存著作物を知っていたかどうかという判例がある
 - 後者の**肯定説の方が有力**ではないかと思われるが、訴訟によって判断が示されないと分からない

- 文化庁が示している依拠性判断の今後の検討事項 (2023年6月19日時点)
 - AI利用者が**既存著作物を認識**し、AIで類似の創作物を生成した場合
 - AI利用者がimg2img等で既存著作物をAIへ入力した場合
 - **特定作者の作品を集中的に学習**した場合とそうでない場合で依拠性判断に違いが出るか

GitHub Copilot利用による著作権侵害



- プログラム著作権特有の問題
 - 非常に小さなルーチンであっても固有の表現があれば著作権は発生
 - ただし、基本的にデッドコピーに準じる程度でないと類似性は認められない
- Copilotの出力が著作権侵害になる流れ
 - CopilotはGitHub.com等から膨大なオープンソースもしくはコード公開プログラムを学習に利用
 - Copyleftライセンスであれば、ライセンスの伝搬と帰属表示の条件
 - 非Copyleftライセンスであれば、ライセンスと帰属表示の条件
 - GitHubは以前に0.1%の確率で**既存コードを逐語的に出力すること**があると認めている
 - Quake IIIのコードをオリジナルコメントを含めたまま出力した例等
 - 逐次的な出力が生成された場合、既存プログラムとの類似性と依拠性が認められる可能性が高い
 - ただし、学習データは複製や翻案の幅広い利用を認めているプログラムがほぼ全て
 - 著作権侵害というより、主にオープンソースの**著作権ライセンス違反の問題に帰結**

ライセンス違反への対応



- ライセンス違反が発生した場合、
 - 非Copyleftライセンスのコード混入であればライセンスと帰属の表示で基本的に解決
 - 業界慣例的には外部指摘後の帰属表示修正で問題ない
 - GPL等のCopyleftライセンスのコード混入の場合、
 - 場合によって派生コード公開が求められる
 - 上記が無理なら、違反による利用許諾の終了
 - 単純使用は継続できるが、基本的に利用するコードの削除が必要
 - GPLに契約性があると既存コード権利者側の判断であれば、
 - ソースコード公開要求、損害賠償請求が可能性として有り得る
 - 損害賠償に関しては、Microsoftが補償プログラムを発表済み
- 対処法：
 - FossilID、Black Duck等のオープンソースライセンススキャナーの導入
 - レビュー、デバッグ工程の強化
 - 既存コードとの同一性フィルタの設定を必ず利用
 - Copilot提案を信用しない

ChatGPT利用による著作権侵害



- プログラムを出力させる場合はCopilotと同様
 - より入出力に制約がないので侵害リスクは大きくなる
 - AutoGPTのような自動化が進めば誰のものでもないコードで氾濫
 - ランセンス管理さえしっかり行えれば局所的に混入事案が発生するCopilotより混乱は少ないかも
- プログラム以外のテキスト作品
 - 一般的なプロンプト入力による出力が既存テキスト著作物と一致、類似するリスクは少ない
 - プログラムと異なり、表現手法の選択幅が格段に大きい
 - ただし、効果的な既存著作物との一致を検出する手法が乏しい
 - 既存テキスト著作物ならある程度は可能
 - 動画著作物を翻案したテキスト作品と看做されるような出力だと現実的には難しい
 - The HEADLINE事案
 - GPTへ追加学習 or プロンプトで他社記事を入力し、多くの出力記事が剽窃と言える内容となった事案
 - 単に参照した記事著作物を複製(再製)する仕組みとなっていたのでは？
- 追加の対処法：
 - 基本はCopilotと同様
 - 特定作者の作品の模倣へ誘導するような入力には極力注意
 - 入力に他者著作物を利用する時は、単なる複製権、翻案権の侵害の可能性を考慮する

イラスト生成AI利用による著作権侵害

- 一般的なテキストプロンプト入力による画像出力の場合、
 - 指示を含んで利用者自身による依拠と学習データによる依拠の二段階
 - 「類似性」の判定は通常著作物と同様
 - 表現上の**本質的な特徴を直接感得**することが条件
- 画風問題
 - 多くのクリエイターが**画風、キャラクターは著作権保護**されると認識している
 - 長年の業界慣習上で熟成された「二次創作」等の独特の著作権解釈を擁する倫理観が存在
 - 実際の所、単なるアイディアに過ぎない**画風は著作権保護の対象外**
 - ネット上に氾濫するAI生成イラストは一部を除き著作権侵害に該当しないとは推測される
 - 絵師とAI術師の間における埋められない認識と情熱のズレが問題を引き起こす
- 週刊「プレイボーイ」AIグラビア事案
 - 生成AI利用で架空人物のグラビア写真集を出版
 - 著作権とパブリシティ権を論争の中心にしたフレームに発展
 - 既存作品との「類似性」がない限りは法的な問題はなかったと考えられる
 - お詫びと販売終了の表明へ追い込まれる



マンション読本事件
大阪地裁平成21年3月26日判決



博士イラスト事件 (東京地裁平成20年7月4日判決)

イラスト生成AI利用による著作権侵害 (続き)

- 既存の画像を入力(img2img等)として使用する場合、
 - 単純な入力画像の加工作業に過ぎないなら、著作権法30条の4は適用されない
 - 出力は、入力画像との「依拠性」が認められるだろう
 - 「類似性」が認められる可能性は高い
 - そもそも既存著作物の翻案権の侵害行為であり、「故意過失」も加わる
 - 従って、既存著作物を利用する場合は許諾を要する
- (狙い撃ち)LoRAによる追加学習のモデルを利用する場合、
 - Stable Diffusion等のモデルに対して追加学習を行う仕組みの一つ
 - 特定作者、キャラクター、画風を模倣する目的とする様々なLoRAモデルが広く頒布
 - 著作権法**30条の4が適用されない解釈**が広まりつつある
 - 一般の基盤モデルの仕組みを考えれば一般的な追加学習は非享受目的と考えられる
 - 立法趣旨、条文的には解釈は難しい意見もあるが、文化庁が出すメッセージは**以前から変化**している
 - 出力は入力画像との「依拠性」が認められる解釈が強い
 - 結局のところ、利用者が特定作品に類似の表現を意図的に出力させる目的なら当然依拠性はある
 - そもそも、利用者が特定作品を知っていればそれだけでも依拠性はある
 - 「類似性」が認められる可能性は比較的高い
 - 画風を模倣するなら必然的に表現が類似する可能性は上がる

まとめ

開発・学習段階

- 一般的にAIモデル開発では無許諾で著作物を利用できる
 - ただし、法域毎に利用用途や市場性等の要因で著作権の権利制限の対象にならないケースも多い
 - 今後は欧州型の権利者側のオプトアウトや学習利用著作物の透明性といったAI規制が世界各地で規定されると推測される

利用・生成段階

- 自律的なAI生成物に著作権は発生しない
 - 法域毎にどのように人間の作業が加われば人間の著作物と判断されるかは若干差異があると考えられる
- AI生成物による著作権侵害の要件は通常と同様
 - 「創作性」のある部分にて「依拠性」と「類似性」が認められれば侵害
 - 依拠性は「利用者+AI」による依拠が認められる

教訓

- 生成AIの出力を過信せず、そのまま利用することは避ける
- 既存著作物を入力のために利用する場合は許諾を取る
- なるべく既存の著作物と一般的な意味での類似がなくなるまで改変する
- AI利用の有無を問わず、作品の制作過程の記録と透明性の確保を心がける

今後

- 各国の法規制議論や判例次第でどこへ向かうか未知数な部分はあるが、アイデア・表現二分論を維持しつつ、グローバルで一致した著作権の緩やかな折り合いがいずれ確立するのだろう
- 非著作物の作品が氾濫する可能性は高いが、人間の著作物との関係でバランスが取られていくのではないかと