NIIオープンハウス2023

生成系AIと法制度
-日EUの状況を中心に-

一橋大学大学院法学研究科ビジネスロー専攻教授
生貝直人　博士（社会情報学）

自己紹介

- 専門分野は情報法・政策、特にデータ・プラットフォーム・AIに関わる日EUの比較法。2012年東京大学大学院学際情報学府博士課程修了、博士（社会情報学）。2012年～2014年情報・システム研究機構新領域融合研究センター特任研究員（国立情報学研究所配属）。科学技術振興機構さきがけ研究員（ビッグデータ基盤領域）等を経て、2021年一橋大学大学院法学研究科ビジネスロー専攻（千代田キャンパス社会人大学院、学術総合センター5～9階）着任。デジタルアーカイブ学会理事等を兼任。

目次

- ①著作権
- ②個人情報・プライバシー
- ③有害情報（誤情報・偽情報・バイアス等）
- ④EU AI規則案
- ⑤学習データの拡大

①著作権

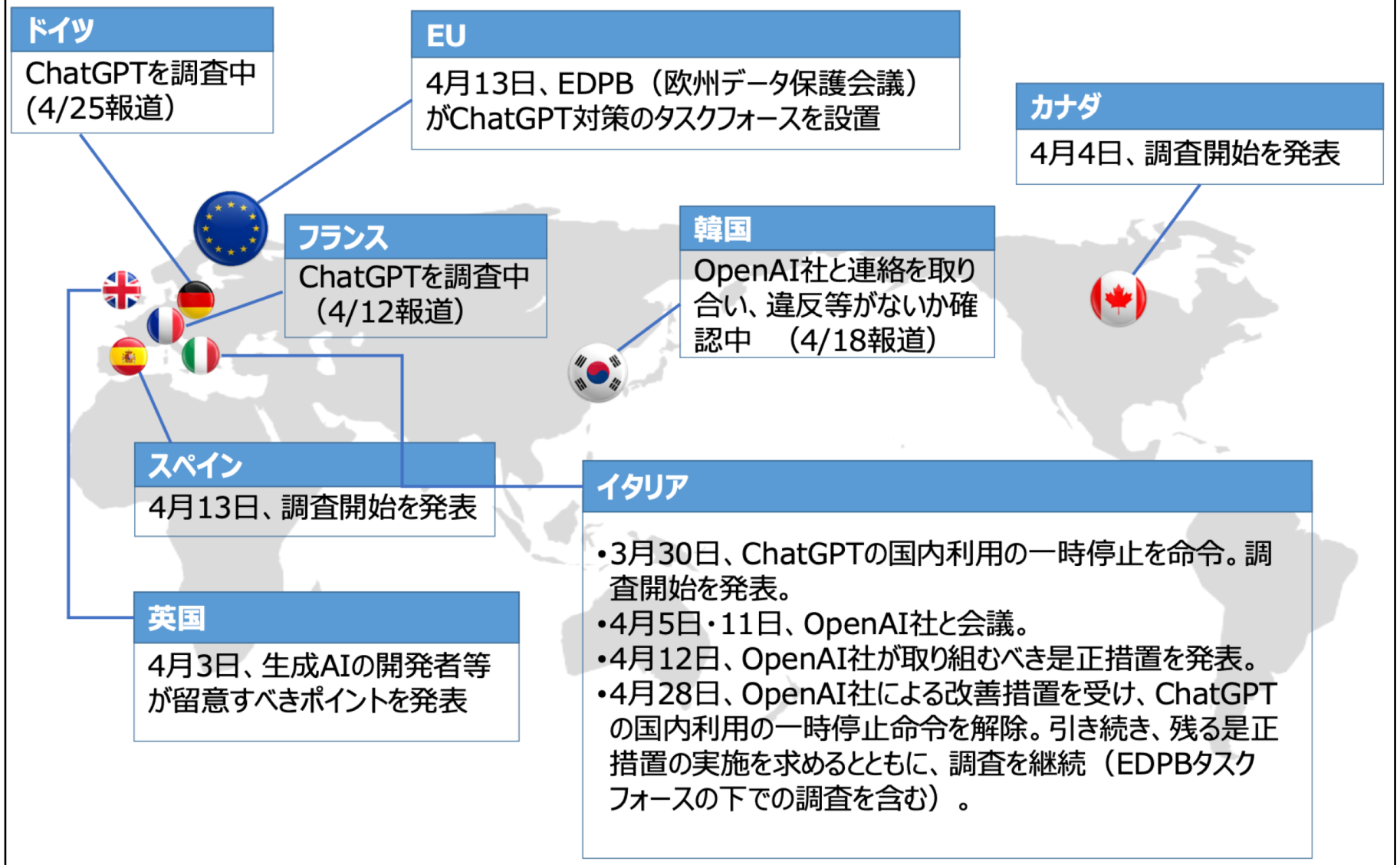
- 開発・学習段階（法域により異なる）
  - 著作権法30条の4により、情報解析等のための「著作物に表現された思想又は感情の享受を目的としない利用行為」は、原則として権利制限の対象。ただし、「**必要と認められる限度**」であることが必要で、「**著作権者の利益を不当に害することとなる場合**」には権利制限の対象とはならない
- 生成・利用段階（およそ各国共通）
  - 既存著作物との「**類似性**」と「**依拠性**」が認められる場合、生成物の利用が著作権侵害となる場合がある
- 生成物の著作権は誰のものか（およそ各国共通）
  - 著作権法が保護するのは人による創作物だけなので、純粋なAI（あるいは人以外の動物等）による作品は保護対象とはならない。ただし**人の創作的寄与が認められる**（AIを道具として利用した）場合には、その人の著作物となる

②著作権：各国状況と論点

- 学習段階（EU・米国）
  - EU：学術研究目的の場合は権利制限の対象、それ以外の営利目的等は権利者のオプトアウトを規定（デジタル単一市場著作権指令3・4条）
  - 米国：フェアユース条項（米国著作権法107条）の解釈

- 若干の論点
  - 「著作権者の利益を不当に害する場合」等を法で細かく決めるべきか
  - 日本法において、例えば営利目的の学習利用からのオプトアウトを認めるべきか。また、権利者への適切な補償を法制化するべきか
  - 特に文章系生成AIの回答が新聞記事などの「クリック」を不要にすることの影響（民主主義のコスト負担）

ChatGPT等に対する各国の個人情報保護当局の対応



内閣府AI戦略チーム第2回資料「諸外国におけるルール等国際動向との関係について」（令和5年5月8日）より

②個人情報・プライバシー

- 指示段階
  - 個人に関する情報のプロンプト入力（目的外利用、第三者提供）
- 出力段階
  - 学習した個人情報の出力可能性
- 学習段階
  - 特に要配慮個人情報の学習利用（取得に際しての本人同意）

- イタリアでのChatGPT提供停止と復帰（GDPR）
  - 学習データからのオプトアウト確保
  - 個人に関する誤った情報の訂正・削除
  - クローリングによる個人情報収集の是非（ウェブ上の大量の顔画像を収集して顔認識機能を操作当局に提供したクリアビューAIとの対比）

OpenAI に対する注意喚起の概要

令和５年６月２日
個人情報保護委員会

当委員会は、令和５年６月１日付けで、OpenAI, L.L.C. 及び OpenAI OpCo, LLC に対し、個人情報の保護に関する法律（平成 15 年法律第 57 号。以下「法」という。）第 147 条の規定に基づき、下記概要のとおり、注意喚起を行った。

なお、本注意喚起は、当委員会が現時点で明確に認識した懸念事項を踏まえたものであり、今後新たな懸念事項を認識した場合には、必要に応じて、追加的な対応を行う可能性がある。

記

1　要配慮個人情報の取得

あらかじめ本人の同意を得ないで、ChatGPT の利用者（以下「利用者」という。）及び利用者以外の者を本人とする要配慮個人情報を取得しないこと（法第 20 条第２項各号に該当する場合を除く。）。

特に、以下の事項を遵守すること。

（１）機械学習のために情報を収集することに関して、以下の４点を実施すること。

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の措置を講じてもな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。

（２）利用者が機械学習に利用されないことを選択してプロンプトに入力した要配慮個人情報について、正当な理由がない限り、取り扱わないこと。

2　利用目的の通知等

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

③有害情報（誤情報・偽情報・バイアス等）

- それ自体違法ではないが所謂「有害（harmful）な情報」について、法はどのように対応しうるか
  - 「ファクトチェック」や「リテラシー向上」
  - SNS等プラットフォーム事業者の「自主的対応」の後押し
- 参考：EUデジタルサービス法（2022年成立）
  - EU域内で4,500万人以上が利用する超巨大プラットフォーム（SNS等）・検索エンジン提供者は、自らのサービスがもたらしうる基本権や民主主義、差別、公衆衛生、青少年保護等への**「システミック・リスク」を評価し、合理的な緩和措置を採り、外部監査を受ける義務**
- 生成AIにおいて同様の法的措置が必要か否か

④EU AI規則案（2021年4月提案時点）

- AIシステムをリスクに応じて4段階に分類した規律を置く
  - 許容できないリスク（禁止されるAI利用行為：潜在意識操作、弱者搾取等）
  - **ハイリスク（適合性評価等の義務）**
  - 限定的リスク（コンテンツがAIで作成されたことの明示等）
  - 低リスク・無リスク（行動規範）
- **ハイリスクAI**に位置付けられる用途（一部）
  - 教育や職業訓練、雇用・労働管理
  - 重要な民間・公共サービス（公的支援金給付、融資、緊急対応措置）
  - 法執行、移民・亡命・国境管理、司法又は民主主義プロセス
- **ハイリスクAI**システム提供者の義務
  - リスクマネジメントシステム構築、適切なデータガバナンス、技術文書、記録保持、透明性と利用者への情報提供、人間による監視、正確性・堅牢性・サイバーセキュリティ

④AI規則案：基盤モデル・生成AI提供者の義務（2023年5月時点）

- 2021年提案当初はChatGPT等の生成AIは想定されておらず、2023年5月の議会修正案では、LLMなど基盤モデル・生成AIに特化したカテゴリーを設けることが提案
  - 基盤モデル（foundation model）：「大規模なデータで学習され、出力の汎用性を考慮して設計され、幅広い特徴的なタスクに適応できるAIモデル」
- 基盤モデル提供者（オープンソースを含む）の義務
  - 健康、安全、基本権、環境、民主主義及び法の支配への**合理的に予見可能なリスク特定と緩和**
  - 適切なデータガバナンス、特にデータに含まれる**バイアスの緩和等**
  - エネルギー使用の削減と効率性向上
  - **川下利用者が本規則を遵守するための情報提供**等の支援（基盤モデル提供側（川上）と利用側（川下）の規制負担を巡る論争）
- 特に生成AI（複雑なテキスト、画像、音声又は映像等のコンテンツを様々なレベルの自律性をもって生成することを特に意図したAIシステム）提供者の義務
  - EU**法違反コンテンツ生成へのセーフガード**を伴う設計・開発
  - **著作権保護学習データの概要公表**（EU著作権指令オプトアウト制度等との組み合わせ）

⑤学習データの拡大

- **法は**（知財法や個人情報保護法のように）データを保護するだけではなく、**活用可能なデータを「増やす」こともできる：データ活用法制**
  - EUオープンデータ指令（2019年成立）：行政保有データのオープン化（G2B）
  - EUデータ法案（2022年提案）：IoT生成データの活用促進（B2B）、民間保有データの政府・公益活用促進（B2G）
  - 欧州ヘルスデータスペース法案（2022年提案）：医療健康データ集約と二次利用
- しかし、どのようなデータを、どこまで、誰に提供すべきか？
  - 例えば医療・健康データ、あるいは国立国会図書館2500万冊のテキストデータ
  - 日本のデータを正しく世界に届けることと「データ主権」のあり方

日本学術会議「研究ＤＸの推進－特にオープンサイエンス、データ利活用推進の視点から－に関する審議について」（2022年12月23日）

- 「EUでは、現在、データガバナンス法やデータ法案、欧州ヘルスデータスペース法案等の、従来の個人情報保護法制や知的財産保護法制のアプローチとは異なるデータ活用の促進に焦点を当てた法制度の整備が進められており、その中には、企業や公的機関が保有するデータを学術研究に利用可能とするための新たな枠組みも含まれる。我が国においても、学術研究分野の個人・非個人情報の保護に関する施策を進めながら、パンデミックへの対応を加速させるための EHR（Electronic Health Record）データの活用を含め、学術研究においてより多くの研究データの活用を可能とするための新たな法的アプローチを検討する必要がある。」（p20）


回答

研究ＤＸの推進－特にオープンサイエンス、データ利活用推進の視点から－に関する審議について

SCIENCE COUNCIL OF JAPAN
日本学術会議

令和４年（２０２２年）１２月２３日

日 本 学 術 会 議