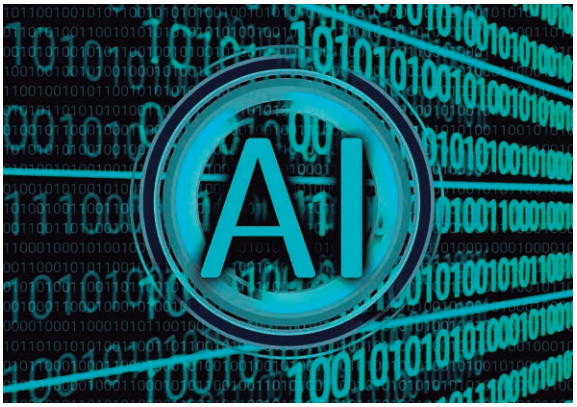
URL https://journal.ntt.co.jp/article/26183
DOI https://doi.org/10.60249/24052001

期待高まる国産生成AI（後編）
―― LLMとAIガバナンス

生成AI（人工知能）の社会実装は，世界的に民主主義やプライバシーといった人権など基本的諸価値・原理への再考を迫られています．大規模言語モデル（LLM：Large Language Model）特集の後編では，法的・倫理的課題の概要とそれに対する国内外の動向について，ユーザ側とサービス・システム提供事業者側の双方から検討します．特にLLM開発事業者のような提供側の企業においては，広く人口に膾炙するほど，企業としての社会的責任を果たすことが求められることから，法的・倫理的配慮を踏まえた「AIガバナンス」の考え方も紹介します．



生成AIの登場で高まる規制議論

前編ではAI（人工知能）の誕生から大規模言語モデル（LLM：Large Language Model）開発動向を概観しましたが，生成AIの利便性に注目が集まる一方で，国内外では，クリエイティブ産業を中心に，危機感の表明とともに，AIに対する規制が求められています．例えば，米国では，AIによるスキル収奪への抵抗として，ハリウッド俳優によるストライキが4カ月にわたって続きました．日本でも，大手出版社が，AI生成画像を使ったグラビア写真集を発売したものの，批判を受け，発売中止となりました．また，TVドラマ『相棒 season22』の最終回では，生成AIによりつくられた主人公の杉下右京のなりすまし動画が全世界に拡散された設定[(1)]でしたが，現実でもこうしたフェイク動画を容易につくることができるようになりました．

加えて，一見してもっともらしい情報や答えを出力するハルシネーション（Hallucination）[*1]があり，偽情報・誤情報の拡散も深刻な問題となっています．

このような背景から，2023年5月に開催されたG7広島サミットの最後に出されたG7広島首脳コミュニケ（同月20日）では，議論すべきテーマとして「著作権を含む知的財産権の保護，透明性の促進，偽情報を含む外国からの情報操作への対応」などが挙げられました[(2)]．

本稿では，以上のような現状を踏まえ，国内外の規制動向を概観した後，日本における生成AIサービスの利活用シーンとLLMの研究開発のシーン，それぞれで生じる法的な問題・議論を紹介します．その後，最近話題になっている「AIガバナンス」の概要を説明し，最後に前後編を通した今後を展望します（本稿は2024年3月14日時点の情報に基づいています）．

国内外の動向の概観

2024年3月13日，欧州議会は，世界初となる，包括的にAIを規制する法案を承認しました．この法案は罰則付きのため，厳しい規制といえます．また，その他の地域においても，**表1**のとおり，生成AIに焦点を当てた規制が始まっています．他方で，日本では，生成AIの登場を踏まえ，これまで複数あったAI関連のガイドラインを統合した，「AI事業者ガイドライン」が作成されています．

生成AIサービスの利活用における法的問題

生成AIサービスを利用してイラストを生成し資料に掲載する，顧客の氏名や生年月日等を入力しリストやグラフを生成させる，あるいは業務効率化のため自組織に生成AIサービスを導入するといった場面では，①知的財産権，②個人情報の保護，③組織内の利用ルールに気を付ける必要があります．

■知的財産権（著作権を中心に）

生成AIの法的問題は，開発・学習段階と利用段階（プロンプトの入力から生成，その後の生成物の利用も含む）の2段階に分かれて検討されることが一般的です．ここでは，利用段階を取り上げます．

まず，著作権との関係では，生成AIにより生成されたイラスト等が，すでに存在している，人によって描かれたイラスト等（既存著作物）に関する権利の侵害となるかが問題となります．これについては，生成AIによる生成物か否かを問わず，「類似性」（他人の著作物と同一か・類似するか）と「依拠性」（他人の著作物をもとに作成されたか）によって判断されることになります[(6)]．これまでも類似性は，裁判の結果の予測がしづらく問題となっていますが（**図1**），生成AIは，開発時に，すでに存在する作品を利用することから，依拠性が特に問題となります．

すなわち，ある作品が生成AI開発時の学習データに含まれているのであれば，その作品に「依拠している」といってよいとも思えるのですが，生成AIに指示を出し

＊1 ハルシネーション：LLMが統計的手法，すなわち，確率論に基づいて文章を生成していることから，結果的に，作成された文章の内容が虚偽となること．

表１　諸外国の生成AI規制の概要

地域	表題	概要	罰則
米国	AI権利章典の青写真	米国国民の権利，機会または重要な資源やサービスへのアクセスに重大な影響を与え得る自動化システムを対象に，５つの原則（①安全で効果的なシステム，②アルゴリズム由来の差別からの保護，③データプライバシー，④ユーザへの通知と説明，⑤人による代替手段，配慮，フォールバック）とこれらに対応する問題点，実践方法・事例をまとめた文書	なし
	AIリスクマネジメントフレームワーク（AI RMF）（国立標準技術研究所（NIST））	NISTサイバーセキュリティフレームワークを参考に，AIに関連するリスクを効果的に管理することを目的としたガイダンス（拘束力なし）．「コア」として，統治，マッピング，測定，管理の４つの対応策を挙げる	なし
	AIの安全，安心，信頼できる開発と利用に関する大統領令（2023年10月）	AI安全保障のための新基準のセクションでは，国家安全保障の観点から，潜在的なデュアルユース大規模基盤モデルを開発する企業等への報告の義務付けや国防および情報機関によるAIの安全かつ倫理的・効果的な使用の保証等により，AIシステムの潜在的リスクから国民を保護する	なし
EU	AI規則	AIの安全規制として，リスクベースで分類し，それぞれに対応した規制を義務付ける．汎用目的AI（顕著な汎用性を示し，広範囲な異なるタスクを有能に遂行することができ，多様な下流のシステムやアプリケーションに統合可能なもの等）の提供者には，①AI規制委員会に通知し，②そのモデルの技術文書の策定・更新，③所定の情報提供義務等，④著作権法を尊重するポリシーの実施，⑤テンプレートに基づき汎用AIモデルの訓練に用いたコンテンツに関する十分に詳細なサマリーの策定・公表等の義務が課される	あり
	AI責任指令案	AIシステムによって引き起こされた損害に対する契約外の民事上の請求に適用．既存の措置に加え，①因果関係の推定による被害者の立証責任の軽減，②上記AI規則における高リスクのAIシステムのプロバイダに対し，関連情報の開示を命じる権限を裁判所に付与する	
英国	AI法案	AI規制庁を設立し，AIに関する当局間調整や検討事項の対応等を所掌させる．また，AIの開発，導入，利用をする企業に対し，AI責任者選任の義務付け（貴族院で審議中）	なし
カナダ	生成AIのガイダンス（ITSAP．00．041）	サイバーセキュリティセンタが発行したガイダンス．生成AIがもたらすリスクを特定し，その軽減するための手段を提示	なし
	生成AIの使用に関するガイドライン	連邦政府機関の職員向けに，連邦機関が生成的AIツールを使用する際の予備的ガイダンスを提供	―
	AI・データ法案	AIの開発事業者に対し，評価，リスク管理，データの匿名化，透明性確保，記録保存の義務を課す（現在，検討中）	あり
中国	生成系人工知能サービス管理暫定規則	生成AIサービス提供事業者に対し，提供前に安全評価の実施，アルゴリズムの事前届出等の義務，提供後に学習に使用されるデータの適法性やユーザの適正使用のモニタリング等の義務を課す	あり
日本	AI事業者ガイドライン	AI開発者，提供者，利用者の遵守事項を記したガイドラインであり，総務省，経産省が作成	なし

出典：各種文献[3]～[5]より情報通信総合研究所作成

た利用者はその作品の存在を知らない可能性もあり，このような場合にまで著作権侵害となるとすると，安心して生成AIを利用できなくなってしまいます．そこで，文化庁の文化審議会著作権分科会法制度小委員会が公表した「AIと著作権に関する考え方について」（以下「考え方」と略します）[8]では，①AI利用者が既存の著作物を認識している場合，②認識していない場合でも，AI学習用データに当該著作物が含まれている場合において，一定の場合に依拠性を認めています（pp.33-35）．

このほか，著作権との関係では，生成物に著作物性が認められるか，すなわち，生

別紙１

別紙２

左を右のイラスト化（赤枠部分）にあたり類似性が否定（なお，赤枠は筆者加筆）．
出典：上記判決（LEX/DB文献番号25449379）別紙１，２[7]より情報通信総合研究所作成

図１　東京地判平成30年３月29日（平成29年（ワ）第672号等）

成AIが生成したイラスト等に「著作権」が認められるかについても，現在，議論と検討が進んでいます．

著作権以外の知的財産について目を向けると，現在，内閣府のAI時代の知的財産権検討会において，著作権以外の産業財産権や不正競争との関係，肖像権・パブリシティ権の問題[9]のほか，声優など実演家の権利に関して，生成動画の実演との関係で権利が及ぶのか[10]，声優の「声」にも権利を認め保護すべきではないかなど，これまで議論されなかった問題が提起されるようになっています[11]．

■個人情報の保護

個人情報保護法は，個人情報取扱事業者（取扱事業者）に対する規制であるため，一般ユーザが生成AIを利用する際には適用されません．他方で，企業が生成AIをユーザとして使う場合は，自社が保有する個人情報の取扱いに注意が必要です．

取扱事業者が注意すべきポイントについては，個人情報保護委員会が2023年6月2日に公表した，「生成AIサービスの利用に関する注意喚起等」[12]が参考になります．この注意喚起の中では，取扱事業者に対して，①生成AIサービスに個人情報を含むプロンプトを入力する場合に，利用目的の範囲内であることを十分に確認すること，②本人の事前同意を得ずに生成AIサービスに個人データを含むプロンプトを入力し，個人データがプロンプトに対する応答結果の出力以外の目的で取り扱われる場合には，個人情報保護法の規定に違反する可能性があるため，このようなプロンプトの入力を行う場合には，当該生成AIサービスの提供事業者が，当該個人データを機械学習に利用しないこと等を十分に確認することが要請されています．

なお，上記注意喚起には，取扱事業者と行政機関等だけでなく，個人情報保護法が適用されない一般ユーザにおける留意点も記載されているため，業務ではAIを使用しないとしても，内容を確認しておく価値があるといえます．

■組織内の利用ルール

企業内で他社が提供する生成AIサービスを活用する場合，組織内のルールを定めることが権利侵害のリスク回避につながります．

利用に際して，重要となるのが，締結される利用契約です．サービスを提供する事業者と交渉し個別に契約をする場合はその契約に，事前に用意されている定型的な利用規約に同意してサービスを利用する場合にはその規約に，それぞれ従うことが原則です．いずれも，特に問題となるのが，①入力したデータの取扱い，②AI生成物に関する権利，利用関係，法的責任（免責条項）の有無です．①については，入力した情報も学習データとして用いることが明記されている場合，前述の個人情報の取扱いに関する問題が生じるほか，企業秘密を入力した場合には，企業秘密の漏洩となる可能性もあります．また，②についても，前述のとおり，第三者の知的財産権を侵害する生成物が生成される可能性があります．この点，AdobeのFireflyのように，権利侵害が生じた場合の補償を定めている場合があります[13]．

組織内ルールの策定にあたっては，日本ディープラーニング協会（JDLA）が，民間企業向けに公表している，「生成AIの利用ガイドライン」[14]が参考になります．また，東京都[15]や福岡県[16]が生成AIのルールを独自に定めており，これらの内容も参考になります．

LLM研究開発における法的問題

LLMの研究開発においては，生成AIサービスの研究開発や利活用とはまた異なった注意が必要です．そこで，以下では，LLM研究開発における法的問題を2つの視点，すなわち，①ユーザの使用により生じ得る権利利益（知的財産，個人情報，偽・誤情報との関係における信用や名誉など）の侵害に対して，研究開発段階でどのような対応が求められるのか，②独占禁止法を中心とする、いわゆる競争法分野におけるLLM研究開発の問題について解説します．

■知的財産関連法規（著作権法を中心に）

著作権者の許諾なく，他人の著作物を複製等する行為は，原則，違法となります．しかし，AI開発のための情報解析のような著作物に表現された思想または感情の享受を目的としない利用行為は，例外的に，著作権者の許諾なく行うことが可能です．ただし，利用態様などから，「著作権者の利益を不当に害することとなる場合」は原則どおり違法になります．このルールによれば，基本的に，著作権者の個別の許諾を得ることなく，他人の著作物を学習用データとして収集・複製し，データセットを学習に利用して，AI（学習済みモデル）を開発することができます．この場合，さらに別途，著作権法47条の5の適用の可能性があります（考え方p.10）．

前述の「考え方」によれば，AIによって生成された著作物に既存の著作物への依拠性が推認された場合，被疑侵害者（著作権を侵害していると疑われている者）の側で依拠性がないことの主張を要するとされています（p.35）．具体的には，AI学習段階において，ログの記録やデータ利用に際しての制約条件の確認（クリアランス）が必要になります．「考え方」では，**表2**のとおり，かなり踏み込んだ記載もされており，法的責任の有無を決する考慮要素が明記されています（プラスは，責任が肯定される方向の要素，マイナスは，責任が否定される方向の要素）．

ただし，利用者側が上記のような主張立証をするには，LLM開発側が，透明性の観点から情報を公開する必要があり，そうでない場合は利用者がリスクを負いながら使用することになります．これに関して，欧州AI法では，LLMの研究開発において，汎用目的AIモデルの訓練に使われたコンテンツについて，フォーマットに基づき十分かつ詳細な概要を記載して公開する義務が課されています．このような規制のない日本において，利用者のリスクにどう配慮するか（あるいは配慮しないか）は，後述するAIガバナンス*2を考えるうえでのヒントになります．

表2　生成AIサービス提供者が侵害主体として責任を負う場合の考慮要素

（+）ある特定の生成AIを用いた場合，侵害物が高頻度で生成される場合 （+）事業者が，生成AIの開発・提供にあたり，当該生成AIが既存の著作物の類似物を生成する蓋然性の高さを認識しているにもかかわらず，当該類似物の生成を抑止する技術的な手段を施していない場合
（−）事業者が，生成AIの開発・提供にあたり，当該生成AIが既存の著作物の類似物を生成することを防止する技術的な手段を施している場合 （−）当該生成AIが，事業者により生成AIの開発・提供にあたり，当該生成AIが既存の著作物の類似物を生成することを防止する技術的な手段を施すなどして，侵害物が高頻度で生成されない場合（仮に，AI利用者が既存の著作物の類似物の生成を意図して生成AIにプロンプト入力するなどの指示を行い，侵害物が生成されたとしても，事業者が侵害主体と評価される可能性は低くなる）

出典：考え方p.37より情報通信総合研究所作成

■個人情報の保護

個人情報保護委員会は，2023年6月2日，「OpenAIに対する注意喚起の概要」と題する文書を公表し，OpenAIに対して行政指導をしました[17]．そこでは，大まかに，①法令が定める場合を除き，要配慮個人情報*3の取得に関し，本人の同意を得ずに取得しないこと（サービス利用者以外も含む），②日本語で利用目的の通知・公表をすべきことを明らかにしました．①については，機械学習による個人情報の収集に関し，なすべき技術的対応が示されています．

なお，プロンプトとして同意を得ていない第三者の個人情報を入力する場合，個人データの提供制限（個人情報保護法27条，海外の場合の同法28条）に抵触する可能性が指摘されています．この点については，前述の個人情報保護委員会の注意喚起の内容からすると，入力した内容について，生成AIサービス提供事業者が学習用データとして利用しないよう対応すれば，原則として個人データの第三者提供とはならないと，現時点では解釈できます[18]．なぜなら，（個人データでなく）個人情報であれば，同法27条の文言上は規制にかからず，仮に個人データとなっていても，同法27条，28条との関係で上記個人情報保護委員会の注意喚起が要求しているのは，「生成AIサービスを提供する事業者が，当該個人データを機械学習に利用しないこと等を十分に確認すること」のみと考えられるからです*4．

海外では，イタリアのデータ保護機関が個人データの取扱いに問題があるとして，2023年3月にChatGPTのサービスが一時停止をして話題になりました[19]．プライバシー保護の国際動向は，GDPR（欧州一般データ保護規則）がデファクトスタンダードとなっており[20]，これは日本企業においてもAIガバナンスを考えるうえで重要です．

■偽情報・誤情報対応

偽情報や誤情報に関する規制は，利活用の際に，違法情報の拡散に当たる場合や名誉毀損・信用毀損など法令に反する行為以外は特に禁止されていません．違法情報以外の偽情報・誤情報対策は，表現の自由との関係で，慎重な議論が必要です．

表3は海外の偽情報規制の概要ですが，規制対象は主に選挙に関する情報です．これは民主主義の根幹にかかわる問題であるためで，実際に，制度・技術的対応が加速するのが選挙時です．DSA（Digital Services Act）については，成立前の2019年のEU議会議員選挙前には，ENISA（欧州連合サイバーセキュリティ機関）が虚偽情報流布活動とその対策について技術的要素と人的要素の両面から分析を行い，フェイクニュース流布・拡散の土壌となっているプラットフォーム事業者に対し，2018年7月までに共通の行動規範を策定して遵守するよう求めました[21]．

また，2020年の米大統領選挙の前には，Microsoftの“Video Authenticator”[22]が公表されました．もっとも，ディープフェイク（DF）検出技術は，いまだ発展途上であり，Metaが2019年に立ち上げたDF競技大会“Deepfake Detection Challenge”[23]では，トップの技術ですら識別精度65%でした．

こうしたDF検出技術等に限界はあるとしても，研究開発段階における対策，例えば，データポイズニング（学習用データを操作・改ざんすることにより，このデータを学習した機械学習モデルを攻撃する手法）や敵対的入力（AIを騙すような入力を行うこと）を防止する対策などは，企業の社会的責任の観点から，求められているといえます．

■競争法分野における問題

近年，GoogleやAmazonなどがデータ取得行為を基軸として市場支配的な地位を有し，それを濫用する行為が問題となっています．同様のことが，LLMに関するビジネスについても懸念されています．

⑴　レイヤ構造と競争環境

生成AI関連ビジネスのレイヤ構造は，大きく3層に大別することができます[24]（**図2**）．1番目は生成AIを支えるインフラです．例えば，クラウドサービスやチッ

＊2　AIガバナンス：本論では，AIの利活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ，そこからもたらされる正のインパクト（便益）を最大化することを目的とする，ステークホルダーによる技術的，組織的，および社会的システムの設計並びに運用を指します（総務省，経済産業省「AI事業者ガイドライン案第1.0版」令和6年3月p.12参照）．

＊3　要配慮個人情報：本人に対する不当な差別や偏見，その他不利益が生じないようにその取扱いに特に配慮を要するものとして政令で定める記述等が含まれる個人情報（個人情報保護法2条3項）．具体的には，人種，信条，社会的身分，病歴，前科，犯罪被害情報，身体・知的・精神障害等，健康診断結果などがあります．

＊4　個人情報と個人データの違い：個人情報とは，生存する個人に関する情報であって，①当該情報に含まれる氏名，生年月日その他の記述等に記載等されることにより特定の個人を識別することができるもの（他の情報と容易に照合することができ，それにより特定の個人を識別することができることとなるものを含む）または②個人識別符号が含まれるものと定義されています．他方，個人データとは，「個人情報」を容易に検索することができるように体系的にまとめた「個人情報データベース等」を構成する個人情報と定義されています．なお，体系的に検索しやすくなっている分，個人データの方が法律上の規制が多くなっています．

表 3　海外における偽情報規制の概要

地域	表題	概要	罰則
米国	保健社会福祉省による啓発策の実施	健康に関する偽・誤情報に関する報告書，対策に向けた専用ページの開設	なし
	ディープフェイク（DF）規制法（州法）	相手方の同意なく，DFを作成，頒布することを禁止（カリフォルニア・テキサス・イリノイ・ニューヨーク等 9 州で同様の規制あり）	あり
EU	デジタルサービス法（DSA）	大規模なオンラインプラットフォーマや検索エンジンに対し，偽情報を含む違法で有害なコンテンツを拡散する際に生じる重大な社会的リスクを特定・分析・査定し，その軽減措置を講ずること等の義務を負う	あり
英国	オンライン安全法	オンラインサービスの提供者に，違法または子どもに有害なコンテンツや活動によるリスクを特定・軽減・管理する義務を課す	あり
カナダ	タスクフォース「SITE」（Security and Intelligence Threats to Elections）	選挙プロセスを妨害する試みを阻止し選挙を保護するため，警察，安全情報局，外務省等が連携する「SITE」が外国からの干渉の脅威に対する監視と評価を強化（一貫した包括的な政策はない）	なし
中国	インターネット情報サービスアルゴリズム推奨管理規定（DF規制法）	音声変換，画像生成や置換などの特定の情報提供サービスの提供者に対し，利用者にその提供元を確認できるようにし，DFによって生成された内容を合理的に示し，大衆に「混同」をもたらさないようにする義務を課す	あり

出典：各種文献より情報通信総合研究所作成

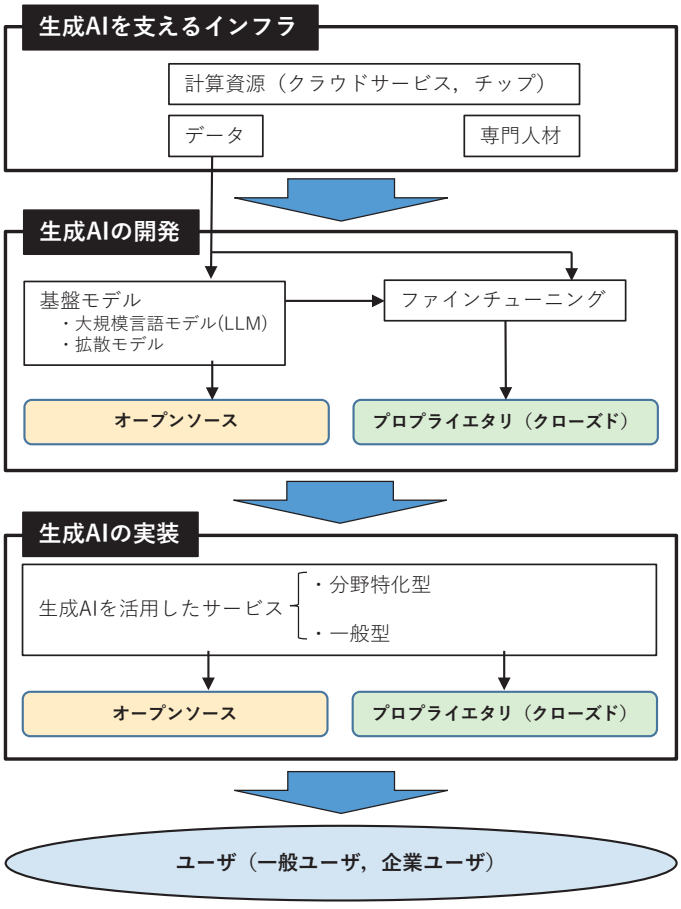


図 2　生成 AI 関連ビジネスのレイヤ構造

プなどの計算資源や，エンジニアや研究者などの専門人材，データなどが該当します．2 番目は生成 AI の開発です．例えば，LLM などの基盤モデルが該当します．またファインチューニングもこのレイヤに分類されます．3 番目は生成 AI の実装です．ここでは具体的にユーザが操作するアプリに生成 AI を実装することが該当します．

以上の各レイヤにおいて，あるいは複数のレイヤにまたがって，競争が生じます．例えば基盤モデルの開発や提供の活発化は，サービス段階での多様なコンテンツの提供を可能とするため，基盤モデル開発段階の競争だけでなく，サービス提供段階の競争も活発化します．

また，特定用途向けにファインチューニングされたモデルを提供する事業形態もあり得ますが，その場合には基盤モデルをどのように調達するかが問題となります．基盤モデルの調達方法には，①資本投入して独自に開発する，②有力な基盤モデル開発事業者と業務提携する，③オープンソースの基盤モデルを活用するなどの方向性が考えられます．

このうち，①は，専門人材やデータセット，計算資源などのインフラが希少であり，それらの利用に関して高額なコストがかかるため，相当程度資本力のある事業者にし

か取り得ない選択肢といえます．他方で，③の方向性は新規参入事業者も選択可能であると考えられます．そのため，オープンソースの基盤モデルの存在は，生成AIの開発・利用に関する参入障壁を下げ，競争を活発化させる重要なファクターになります．

⑵　想定される競争と反競争的行為

以上の競争の機会や特徴を前提とした場合，競争法違反となる可能性がある事業者の行為や問題となる場面がいくつか想定できます．第1に，基盤モデルやデータセットへのアクセスについてです．例えば，基盤モデルの調達方法として上記②のように有力な基盤モデル開発事業者と業務提携する場合を想定してみましょう．この場合に，基盤モデルを提供し，かつ生成AIを活用したサービスを提供する事業者が，新規参入者や競合他社による当該基盤モデルへのアクセスを制限すると，ユーザへのサービス提供にかかる競争の機会が損なわれることになります．このような事象は，基盤モデルの事前学習やファインチューニングの際に必須となるデータセットについても同様に起こり得ます．

第2に，自己優遇です．基盤モデル提供事業者が，自社の商品・サービスが他社のものよりも有利に出現するように基盤モデルを開発したり，基盤モデルを利用したサービス提供事業者が当該サービス提供時に自社商品・サービスを優遇したりすることがあり得ます．自己優遇が競争法違反となる理由については，講学上も議論の途上にありますが，イノベーション阻害などの悪影響が生じ得るとされています[25]．少なくとも近年では巨大デジタル・プラットフォーマによる類似行為は規制対象とされつつあり，EUでは競争法違反を認めた事例も登場しています[26]．

第3に，抱き合わせや囲い込みです．例えば，クラウドサービス市場で有力な地位にある事業者が同サービスの提供条件として，自社製生成AIの基盤モデルの使用を抱き合せる場合などが想定されます．この場合には基盤モデルの製造・提供段階での競争が阻害される可能性があります．

このほかにも，人材獲得競争との関係で，事業者間で引き抜き防止や賃金水準について合意することがカルテルに該当するかという問題や，合併等のかたちで異なるレイヤの事業者の統合が企業結合規制に抵触するかという問題も想定されます．

企業におけるAIガバナンス

■AIガバナンスとは

現在，日本にはAIを直接規制する法制度はなく，ソフトローアプローチを採用するといわれています．もっとも，個別法令では，AI技術の進展に伴い，法改正を伴うハードローで規制している側面もあります．企業は，これまで取り上げた法令に限らず，倫理的な配慮から，業界のガイドラインなどソフトローに配慮することが必要です．また，経済産業省等が策定したAI事業者ガイドライン案[27]では，企業に，ステークホルダーの一員として，AIガバナンスの構築が求められています．ここでは，AIガバナンスについて，「AIの利活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ，そこからもたらされる正のインパクト（便益）を最大化することを目的とする，ステークホルダーによる技術的，組織的，及び社会的システ

表4　「高度なAIシステムを開発する組織向けの広島プロセス国際行動規範」の概要

	遵守すべき行動規範	具体例
1	リスクの特定，評価，軽減のための適切な措置	レッドチーミング等の評価方法の組合せや多様な内部テスト手段や独立した外部テスト手段の採用
2	導入後の脆弱性・インシデント等への対応	インセンティブ型の脆弱性報告制度
3	アカウンタビリティ向上	重要な新規公表事項全てを含む透明性報告書の公表義務
4	ステークホルダーとの情報共有とインシデント報告	AIのライフサイクル全体にわたる他組織との協力と関連情報の共有と公表，公的機関との連携・情報共有
5	AIガバナンスとリスク管理方針の策定	リスク管理のガバナンス方針の策定・開示，実施のための組織的メカニズムの導入
6	セキュリティ対策	アクセス制御，リスク評価，技術的・制度的解決策の実施や脆弱性管理プロセスの導入
7	電子透かしやコンテンツ認証，来歴メカニズム開発・導入	ユーザがAI生成物と判断できるツールやAPIの開発や当該技術への研究開発投資，ラベリングや免責事項の表示
8	リスク軽減策への研究とその投資	社会的弱者や知的財産権やプライバシーの保護，偽・誤情報への対策のための研究とその投資等
9	重大な社会課題を解決するAIの開発の優先	SDGsを支援し，グローバルな利益をもたらすAI開発をし，市民社会やコミュニティ・グループとの協力とし，優先課題を特定し，重大な社会課題（特に気候危機等）の解決策を開発
10	国際的技術規格の開発推進	行動規範7の技術に関し，国際的技術標準の開発をめざす
11	個人データ・知的財産の保護	データの質の管理をするための適切な措置を講じる

出典：上記文書より情報通信総合研究所作成

ムの設計及び運用」であると定義されています．

■広島AIプロセスから見る企業に求められるAIガバナンス

「AI事業者ガイドライン案」はこれまであったガイドラインを統合・見直したものです．これにより，AIガバナンスは，原則から実践へ向け，より具体化したといえます．この傾向は，2023年の広島AIプロセスにおいても同様であり，例えば，広島AIプロセスの成果文書の1つである「高度なAIシステムを開発する組織向けの広島プロセス国際行動規範」では，これまでよりも企業や公的機関に求められるものが具体化されています（**表4**）．

展　望

本稿前編のとおり，生成AIは，①スケーリング則の発見，自己教師あり学習などの技術的要因と，②研究開発への大規模投資といった経済的要因によって，言語モデルの大規模化が加速し，特に2020年代以降，競争が激化しました．そして，生成AIの急速な普及の背景には，利便性もさることながら，③RLHF（Reinforcement Learning from Human Feedback：人間のフィードバックからの強化学習）などにより出力制御を行うといった，開発側の法的・倫理的な配慮があります．ただし，本稿後編で述べてきたように，法的・倫理的課題への対応はいまだ発展途上であり，それを規制するルールもまた同様です．それゆえに，規制面では，今まさに世界中で議論されているAIガバナンスの議論が重要になります．

他方で，研究開発の技術的側面においては，パラメータ数千億超の大規模モデルの開発は計算量の膨大化や電力の大量消費を伴い，サステナブルでないことが指摘されていることから，超軽量モデルの需要増が見込まれ，小規模モデルの高性能化が急務であるといえます．小規模モデルの研究開発は，モデルの大規模化と投資規模の拡大による一部の企業の寡占という懸念を回避することにもつながります．

今後，小規模モデルが複数登場し普及した場合には，それらの相互運用性（オープン性）の確保やその一手段である標準化が重要になります[28]．これまで，標準化は，世界市場のシェアを獲得し，自国の規格を世界の標準規格として普及させるという技術覇権競争としてとらえられることが多かったのですが，AI技術の分野では，むしろ国際協調が強調されています．これは，日本企業においても，小規模モデルの研究開発の際に市場を拡大できるメリットとなり得る一方，その利益を十分に享受するためには，著作権や個人データの保護に関する国際的な制度的調和を含めた相互運用性のあるAIガバナンスの構築という課題を乗り越えることが必要になります．

■参考文献

(1) https://post.tv-asahi.co.jp/post-245124/
(2) https://www.soumu.go.jp/menu_news/s-news/01tsushin06_02000277.html
(3) 新保：“AI規制の国際動向，”都市問題，Vol. 115, No. 2, p. 18，2024.
(4) 松尾：“成立間近のEU「AI法」で留意すべきAI利用者への影響，”週刊金融財政事情，No. 3532，p. 34，2024.
(5) 栗原：“カナダ，欧米におけるAI規制法案の動向からみるAIガバナンス，” InfoCom T&S World Trend Report，No. 401，p. 20，2022.
(6) https://www.bunka.go.jp/seisaku/chosakuken/pdf/93903601_01.pdf
(7) 判例データベース：LEX/DBインターネット，文献番号 25449379.
(8) https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/bunkakai/69/pdf/94022801_01.pdf
(9) https://www.kantei.go.jp/jp/singi/titeki2/ai_kentoukai/gijisidai/dai4/index.html
(10) 栗原：“メタバースを中心とするバーチャルリアリティにおける著作権法の「実演」に関する一考察，” 情報通信政策研究，Vol. 6，No. 2，p.15，2022.
(11) 荒岡・篠田・藤村・,成原：“声の人格権に関する検討，”情報ネットワーク・ローレビュー，Vol. 22，p. 24，2023.
(12) https://www.ppc.go.jp/files/pdf/230602_alert_generative_AI_service.pdf
(13) 岡田・羽深・佐久間：“連載AIガバナンス相談室　第2回AIガバナンス「AI利用事業者編」，”ビジネス法務，Vol. 24，No.3，p. 66，2024.
(14) https://aismiley.co.jp/ai_news/jdla-chatgpt-llm-generativeai/
(15) https://www.metro.tokyo.lg.jp/tosei/hodohappyo/press/2023/08/23/14.html
(16) https://www.pref.fukuoka.lg.jp/press-release/generative-ai-fukuoka-guideline.html
(17) https://www.ppc.go.jp/files/pdf/230602_alert_AI_utilize.pdf
(18) 岡田・堺：“文書要約または文書作成に関する社内ルールの整備，”ビジネス法務，Vol. 23，No. 11，p. 25，2023.
(19) https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english
(20) https://yuhikaku.com/articles/-/18653
(21) 湯淺：“EUにおけるフェイクニュース対策，”日本セキュリティ・マネジメント学会誌，Vol. 32，No. 3，p. 45，2019.
(22) https://blogs.microsoft.com/on-the-issues/2020/09/01/disinformation-deepfakes-newsguard-video-authenticator/
(23) https://www.kaggle.com/c/deepfake-detection-challenge
(24) https://assets.publishing.service.gov.uk/media/65081d3aa41cc300145612c0/Full_report_.pdf
(25) 林：“デジタル・プラットフォーム事業者による自己優遇行為と反トラスト法，”法律時報，Vol. 94，No. 8，p.75，2022.
(26) 千葉：“デジタル化社会の進展と法のデザイン，”商事法務，p. 246，2023.
(27) https://www.soumu.go.jp/main_content/000935246.pdf
(28) https://www.meti.go.jp/meti_lib/report/2022FY/000802.pdf


株式会社情報通信総合研究所
研究員　酒井基樹（写真左）
研究員　成冨守登（写真なし）
主任研究員　栗原佑介（写真右）