CITY-YUWA NEWSLETTER Vol.40

CITY-YUWA PARTNERS

2024.5.30

# EU AI Act の概要

## I. はじめに

一昨年秋の生成 AI の登場以降、AI の活用場面が飛躍的に拡大するとともに、知的財産権の侵害、偽情報の生成・発信等、従来の AI にはなかったような新たなリスクが懸念される中、AI の利用に伴うリスクを管理しつつその便益を最大化するための社会的取り組み(AI ガバナンス)についての議論が世界各国で進められてきました。かかる議論の成果として、我が国においては、本年 4 月 19 日に AI の開発・提供・利用を行う者を対象にそれぞれに必要な取り組みの指針を定めた AI 事業者ガイドライン(第 1.0 版)[1]が公表されました。また、EU においては、本年 3 月 13 日に、AI の提供者、利用者等の法的義務を定めた包括的 AI 規制としての Artificial Intelligence Act [2](以下「AI Act」といいます。)が欧州議会において承認され、本年 5 月 21 日に EU 理事会において承認されるに至りました。

AI 事業者ガイドラインは、AI の開発・提供・利用を行う各主体が取り組むべき事項の指針として、①人間中心、②安全性、③公平性、④プライバシー保護、⑤セキュリティ確保、⑥透明性、⑦アカウンタビリティを掲げています。他方 AI Act も、後記のとおり①人間による主体性の確保と監督、②技術的堅牢性と安全性、③プライバシー保護とデータガバナンス、④透明性、⑤多様性、非差別と公正さ、⑥社会と環境への福祉的配慮、⑦説明責任の 7 つの原則をその行動規範の基礎と位置付けています。このように、AI 事業者ガイドラインも AI Act もその基礎とする理念はほぼ重なるものとなっています。

他方、AI 事業者ガイドラインは、法的拘束力のないソフトローであるのに対し、AI Act は法的拘束力を有する規範であるという点で大きな違いがあります。AI 事業者ガイドラインは、「AI 開発・提供・利用にあたって必要な取組みについての基本的な考え方を示すもの」であり、AI 活用に取り組む全ての事業者が、「本ガイドラインを参考の一つとしながら」、「自主的に具体的な取組を推進することが重要」としています。これに対して、AI Act においては、AI のプロバイダー、利用者等の法的義務が詳細に規定され、かかる義務違反には罰則が課され得るものとなっています。その意味で、AI Act は、AI 事業者ガイドラインと同様の方向性を目指しつつも、ガバナンスの方法としては AI 事業者ガイドラインの対極にあるものと言えます。

国内の事業者にとっては、当面 AI 事業者ガイドラインへの対応が主要な関心事となると思われますが、AI Act についても、EU 域内において事業展開を行う事業者に直接の影響が及ぶことはもとより、我が国における AI ガバナンスの今後の方向性に大きな影響を及ぼすものとして、その内容及び運用の動向についてフォローしておくことは極めて有益なことと思われます。

そこで、本ニューズレターにおいては、AI Act の制定の経緯及び基礎となる理念を背景として概説したうえで、AI Act の主要な条項についてその概要をご紹介します。

## II. AI Act 制定の背景

### 1. AI Act 制定に至る経緯

| | |
|---|---|
| 2018 年 3 月: | 欧州委員会の科学新技術倫理グループ(EGE)が「AI・ロボティクス・自律システムに関する宣言」を公表。 |
| 2018 年 4 月: | 欧州委員会が AI 戦略「欧州の AI」を公表。 |
| 2019 年 4 月: | 欧州委員会の設立した AI ハイレベル専門家グループ(HLEG)が「信頼できる AI のための倫理ガイドライン」(以下「AI 倫理ガイドライン」という。)を公表。 |
| 2020 年 2 月: | 欧州委員会が「AI に関するホワイトペーパー:優越と信頼に向けた欧州アプローチ」(以下「AI ホワイトペーパー」という。)を公表。 |
| 2020 年 10 月: | 欧州議会が「AI・ロボティクス・関連技術の倫理フレームワークに関する欧州委員会への勧告」を決議。 |
| 2021 年 4 月: | 欧州委員会が「AI Act(Artificial Intelligence Act)案」を公表。 |

[1] AI 事業者ガイドラインは、従来の総務省主導の「国際的な議論のための AI 開発ガイドライン案」と「AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～」及び経済産業省主導の「AI 原則実践のための ガバナンス・ガイドライン Ver. 1.1」を統合・見直して策定されました。

[2] 本ニューズレターにおける説明の対象とする AI Act は、かかる本年 3 月 13 日の欧州議会で承認されたバージョンの規制案(https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html)に基づいております。

2022 年 12 月: EU 理事会が AI Act 案の修正案を採択。
2023 年 6 月: 欧州議会が AI Act 案の修正案を採択。
2023 年 12 月: EU 理事会と欧州議会とが AI Act 案の修正案につき暫定的な合意に到達。
2024 年 3 月 13 日: 欧州議会が欧州議会修正案と EU 理事会修正案の合意内容を harmonized rule として盛り込んだ AI Act 案を承認。
2024 年 5 月 21 日: EU 理事会が AI Act 案を承認。

## 2. AI ホワイトペーパーが示した AI が基本権に及ぼすリスクについての考え方

欧州委員会の AI ホワイトペーパーは、AI の基本権にもたらすリスクについて、大要以下のように述べています。

① AI の利用は、基本権(表現の自由、集会の自由、人間の尊厳、性別・人種・民族・宗教・信条・身体的障害・年齢・性的指向等により差別を受けない権利、司法的救済を受ける権利及び消費者として保護を受ける権利)に悪影響を及ぼす可能性がある。
② これらのリスクは、AI システムの設計(人的監視も含む)自体から生じる場合もあれば、バイアスを是正しないデータを利用することにより生じる場合もある。
AI はこれまで人間により行われてきたことを人間に代わって行うことができる。それにより、人は、AI によるあるいは AI の助けを得て行われる行為の影響を受けまたそのようにしてなされる決定に服するようになるが、その場合、かかる行為や決定を理解し必要に応じ異議を申し立てることが難しくなる可能性がある。
③ AI は、人々の日常生活を追跡し分析することを可能にする。例えば、AI が、データ保護法に反する形で、公的機関により大規模監視に用いられたり、雇用者により被用者の行動を監視することに用いられる可能性がある。また、AI が、匿名化された大規模データを分析しその相互関係を把握して個人データとして再識別することに用いられ、それにより新たな個人データ保護の侵害のリスクを生ぜしめる可能性がある。
④ AI は、オンライン仲介業者がユーザーに提供する情報に優劣を設けコンテンツの内容を調整することに用いられる。選択的に提供されるデータの内容やそのアプリケーションによって、表現の自由、個人データ保護、プライバシー、政治的自由に影響が生じる可能性がある。
⑤ 人間による意思決定もバイアスから逃れることはできないが、AI によりバイアスが提供される場合、その影響はより深刻である。その場合、人間の行為を支配するために設けられている様々な社会的コントロールを受けることなく多くの人々が差別され悪影響を受けることとなるからである。
⑥ 多くの AI が有する特定の性質、特に、ブラックボックス効果、複雑性、予測困難性、部分的な自律的行動は、基本権を保障するための現行の EU 法の効果的執行を困難にする可能性がある。執行機関や関連当事者は、ある決定に AI が介在する場合、かかる決定がどのように形成されたかを疎明することが難しくなり、その結果、当該決定に係るルールが尊重されなくなる可能性がある。かかる決定が当事者に不利な影響を及ぼすものである場合、当該当事者がかかる決定について司法の審査を求めることが困難になる可能性もある。

これらのリスクの考慮が AI Act の策定にあたっても重要な指針となっていたことが伺われます。

## 3. AI 倫理ガイドラインが示した AI の信頼性確保のための 7 つの原則

AI HLEG が公表した AI 倫理ガイドラインは、AI の信頼性と倫理的健全性の確保に資することを目的とした 7 つの倫理原則を策定しました。
この 7 つの原則は、①人間による主体性の確保と監督、②技術的堅牢性と安全性、③プライバシー保護とデータガバナンス、④透明性、⑤多様性、非差別と公正さ、⑥社会と環境への福祉的配慮、⑦説明責任からなっています。
かかる倫理原則を受けて、AI Act 前文(27)は、かかる原則について以下のとおり説明した上で、これらの原則が AI Act における行動規範を策定するにあたっての基礎となるべきであると述べています。

① 人間による主体性の確保と監視とは、AI システムが人間のために役立つツールとして開発・使用され、人間の尊厳と個人の自律性を尊重し、人間が適切に制御・監督できる形で機能することを意味する。
② 技術的堅牢性と安全性とは、問題が発生した場合の堅牢性及び、第三者による違法な使用を可能にする

AI システムの使用や性能の変更の試みに対する回復力を有し、意図しない危害を最小限に抑えることができるように AI システムが開発され使用されることを意味する。

③ プライバシー保護とデータガバナンスとは、AI システムがプライバシーとデータ保護の規則に従って開発・使用され、同時に、品質と完全性の面で高い基準を満たす形でデータを処理することを意味する。

④ 透明性とは、AI システムが適切な追跡可能性と説明可能性を実現するように開発され使用されることを意味する。そして、人に AI システムと通信又は交信していることを認識させ、AI システムの利用者には当該 AI システムの能力と限界について、また当該 AI システムにより影響を受ける人にはそのことに関していかなる権利を有しているかを、それぞれ知らしめることを意味する。

⑤ 多様性、非差別と公正さとは、AI システムが、多様な主体を包含し、平等なアクセス、男女平等、文化的多様性を促進するような形で開発・使用されるとともに、連邦法や国内法で禁止されている差別的インパクトや不当な偏見が生じないようにすることを意味する。

⑥ 社会と環境への福祉的配慮とは、個人、社会、民主主義への長期的な影響を監視・評価しながら、AI システムが持続可能で環境に優しく、すべての人間に恩恵をもたらす形で開発・使用されることを意味する。

## III. AI Act の主要な規定

### 1. 一般規定（Chapter I）

(1) 定義(Article 3)

Article 3 では、多くの用語について定義規定が設けられていますが、AI Act の概要の理解のために最低限必要と思われる用語の定義を以下に掲げました。

AI システム(Article 3(1)):様々なレベルの自律性で動作するように設計された機械ベースのシステムであって、導入後に適応力を示すことがあり、かつ明示的又は黙示的な目的のために、提供を受けたインプットから、物理的又は仮想的な環境に影響を与え得る予測、コンテンツ、推奨又は決定などのアウトプットを生成する方法を推測するもの。

プロバイダー(Article 3(3)):有償・無償を問わず、AI システム又は汎用 AI モデルを開発する又は開発された AI システム又は汎用 AI モデルを保有して自己の名称若しくは商標の下にそれを市場に投入し若しくはサービスに組み込む自然人若しくは法人、官公庁、機関又はその他の団体。

輸入者(Article 3(6)):第三国の自然人又は当該国で設立された法人の氏名・名称若しくは商標を載せている AI システムを市場に投入する、EU 域内に所在し又は EU 域内で設立された自然人又は法人。

販売業者(Article 3(7):EU 域内の市場において AI システムを利用できるようにするサプライチェーン内の自然人又は法人であって、プロバイダー又は輸入者以外の者。

利用者(Article 3(4)):自らの権限の下で AI システムを使用する自然人又は法人、公的機関、代理店又はその他の機関をいう。ただし、AI システムが個人的な非職業的活動の過程で使用される場合を除く。

(2) 適用範囲(Article 2 Paragraph 1)

Article 2 Paragraph 1において、AI Actは以下について適用される旨規定されています。

(a) AIシステムをEU域内で市場に投入する若しくはサービスに組み込む又は汎用AIモデルを市場に投入するプロバイダー。設立場所がEU域内か第三国かは問わない。
(b) AIシステムの利用者であって、EU域内で設立された又はEU域内に所在する者。
(c) アウトプットがEU域内で利用される場合、AIシステムのプロバイダー及び利用者であって、設立地又は所在地が第三国の又は第三国に所在する者。
(d) AIシステムの輸入者及び販売者。
(e) 製造業者であって、AIシステムを自らの製品とともに自らの名称又は商標を以ってEU域内で市場に投入する若しくはサービスに組み込む者。

(f)　EU域外で設立されたプロバイダーのEU域内代理人。
(g)　EU域内に所在して影響を受ける者。

## 2. 禁止されるAIシステム（Chapter Ⅱ）

AI Act は、リスクベースのアプローチを取っており、リスクに応じて本項及び次項に記載されるカテゴリーを設けてそれぞれ規制を定めています。本項に記載する「禁止されるAIシステム(Prohibited AI Practices)」というカテゴリーについては、リスクが許容できる限度を超えていることからそもそも禁止されるものとしています。かかるカテゴリーに含まれるAIシステムとして以下のものが定められています。

(1) サブリミナル技術を使用するAIシステム
人に身体的又は心理的な損害を与える又は与え得る方法で人の意識を超えたサブリミナル技術又は人の行動を実質的に歪める若しくは人々の情報に基づく意思決定を損なう操作的・欺罔的な技術を導入し、その影響がなければ取らなかったであろう行動を取らせることとなるAIシステム。

(2) 子供や障害者等を搾取するAIシステム
人に身体的又は心理的な損害を与える又は与え得る方法で人の行動を実質的に歪めるために年齢、身体的又は精神的障害に起因する特定の集団の脆弱性を利用するAIシステム。

(3) ソーシャルスコアリング
社会的行動又は既知若しくは予測される個人的若しくは人格的特徴に基づいて一定期間にわたって自然人の信頼性をソーシャルスコアで評価又は分類するための AI システムであって、以下のいずれか又は両方をもたらすもの。
①データが元々生成又は収集された文脈とは無関係な社会的文脈において、特定の自然人又はその集団全体を不利に扱うこと。
②特定の自然人又はその集団全体に対して、その社会的行動又はその重大性に照らして不当又は不相当に不利な扱いをすること。

(4) 自然人のプロファイリングやパーソナリティや性格の評価のみに基づき当該自然人が犯罪を犯す可能性を予測するAIシステム。

(5) インターネットや監視カメラ映像からの顔情報の無差別収集による顔認識データベースを作成する AI システム。

(6) 職場、教育機関における感情認識システム。

(7) 機微な特徴(例:人種、市民権、宗教、政治的指向等)を推測するため生体データに基づき自然人を分類する生体分類システム。

(8) 公衆がアクセス可能な空間における法執行目的でのリアルタイム遠隔生体認証システムの利用。ただし、以下の目的のために厳密に必要な場合は除く。
①行方不明の子供を含む、特定の潜在的な犯罪被害者の的を絞った捜索
②自然人の生命又は身体の安全に対する特定の、重大かつ差し迫った脅威又はテロ攻撃を防止すること
③少なくとも 4 年の最長期間の拘留刑又は拘留命令により罰せられる犯罪の実行者又は被疑者の発見、位置特定、特定又は起訴

## 3. ハイリスクAIシステム（Chapter Ⅲ）

本項に記載するハイリスクAIシステムのカテゴリーに属するAIシステムは、禁止されるものではないものの、以下に示すような要件、義務等に関する規定の適用を受けます。

(1) ハイリスク AI システムに含まれる AI システム(Article 6)

(a) 以下の条件の両方を満たす AI システム。
①AI システムが、Annex Ⅰ に記載されている法令の対象となっている製品の安全部品として使用されることを意図しているか、又はそれ自体が製品であること。
②AI システムを安全部品とする製品又は製品としての AI システム自体が、Annex Ⅰ に掲げる法令に基づく第三者による適合性評価義務の対象であること。

(b) AnnexⅢに定める用途に用いられる AI システム(自然人の意思決定に重大な影響を与えないなど、自然人の健康・安全・基本権に重大なリスクとならない場合として、Article 6 paragraph 3 に列挙された場合にあたる場合を除く。)

ア 生体認証
① 遠隔生体認証用 AI システム
② 生体的特徴による分類を目的とする AI システム
③ 感情認識を目的とする AI システム
イ 重要インフラ
重要なデジタルインフラ、道路交通、水・ガス・暖房・電気の供給の管理・運営における安全部品用 AI システム
ウ 教育・職業訓練
① 教育機関及び職業訓練機関へのアクセス又は割当て決定
② 学生の成績評価
③ 入試評価
④ 学生の禁止行為の探知・監視
エ 雇用、労働者管理及び自営業へのアクセス
① 採用・選考
② 雇用関係の条件決定、昇進、雇用関係の終了、タスク割り当て、パフォーマンス評価
オ 必須の民間・公共サービスへのアクセス
① 公的支援給付及びサービスの利用資格の評価、付与、減額、取消し又は再請求
② クレジットスコア
③ 生命保険・医療保険に関するリスク評価
④ 緊急時初動対応サービスの派遣(消防士、医療等)に関する分類、優先順位の決定等
カ 法執行(主体は法執行機関及びその補助機関)
① 自然人が犯罪被害者となるリスクの評価
② ポリグラフ等
③ 証拠の信頼性評価
④ プロファイリングや性格特性又は過去の犯罪行動の評価に基づく犯罪発生の予測
⑤ 犯罪捜査・訴追過程でのプロファイリング
キ 移民、亡命、国境管理(主体は所管の公的機関)
① ポリグラフ等
② 入国者の安全保障上のリスク、不正移民のリスク又は健康上のリスクなどの評価
③ 亡命、査証及び滞在許可申請の審査
④ 移民、亡命、国境管理の関連での自然人の特定
ク 司法及び民主的プロセスの運営
① 司法当局及びその補助機関による事実と法律の調査・解釈及び法律の適用の支援並びに代替的紛争処理手続きにおける同様な利用
② 選挙や住民・国民投票の結果や投票行動に影響を与える利用

(2) ハイリスク AI システムが満たすべき要件

(a) リスク管理システムの構築、実施等(Article 9)

リスク管理システムは、以下のステップから構成されることが求められる。
① 本来の用途に従って利用された場合に生じ得る健康、安全、基本権に対するリスクの特定と分析
② 本来の用途に従って利用される場合又は合理的に予見可能な誤用が行われた場合にリスクの評価
③ 市販後モニタリングによるリスク評価
④ 適切なリスク管理措置の採用

(b) データ及びデータガバナンス(Article 10)
データを用いて学習するAIモデルを採用しているハイリスクAIシステムは、Article 10のparagraph2~5に定める一定の品質基準を満たす学習・検証・テスト用データセットを用いて開発されなければならない。

(c) 技術文書(Article 11)
Article 11に定める要件を満たす技術文書を上市前又はサービス提供前に作成し、最新の状態に維持しなければならない。

(d) 記録保存(Article 12)
システム動作中の自動ログ機能を備えなければならない。

(e) 透明性及び利用者への情報提供(Article 13)
① 利用者がAIシステムのアウトプットを解釈して適切に利用できるよう、透明性を確保しなければならない。
② 利用者に関連するアクセス可能で理解可能な、簡潔、完全、正確かつ明確な情報が含まれる使用説明書を添付しなければならない。
③ 使用説明書には、プロバイダーの身元及び連絡先、AIシステムの特徴、能力及び性能の限界に関する事項(本来の用途、正確性・堅牢性・サイバーセキュリティのレベル、健康及び安全又は基本的権利に対するリスクにつながる可能性等)、適合性評価時点からの性能の変更点、人的監視措置、必要なハードウェア、想定寿命及び適切な機能を確保するために必要な保守・ケア等が記載されなければならない。

(f) 人的監視措置(Article 14)
ハイリスクAIシステムは、適切な人間・機械間インターフェースツールを含め、AIシステムが使用されている期間中、人間が効果的に監督できるような方法で設計・開発されなければならない。
人間による監視は、プロバイダーが特定した上で、自らAIシステムに組み込むか、利用者による実施に適した措置を通じて確保する。
人間による監視措置は、監視業務要員が、AIシステムの能力及び限界を十分に理解し、その動作を適切に監視し、異常、機能不全及び予想外の性能の兆候をできるだけ早く検知して対処すること等により利用者に提供されなければならない。

(g) 正確性・堅牢性・サイバーセキュリティ(Article 15)
ハイリスクAIシステムは、適切なレベルの正確性、堅牢性、サイバーセキュリティを達成し、ライフサイクルを通じて一貫した性能を発揮するように設計・開発されなければならない。

(3) ハイリスクAIシステムに関するプロバイダーの義務(Article 16)

(a) ハイリスクAIシステムが本規則の要件に準拠していることを確保すること。
(b) プロバイダーの名称、連絡先の表示。
(c) 品質管理システムを策定・文書化すること。
(d) 技術文書を作成すること。
(e) 自己の管理下にある場合、自動的に生成するログを保管すること。
(f) 上市前又はサービス提供前に関連する適合性評価手続を受けること。
(g) EU適合性宣言の作成及びCEマークの貼付。
(h) EUデータベースへの登録。
(i) 必要な是正措置と情報提供。
(j) 当局への協力:要請があった場合、ログへのアクセス提供、本規則の要件遵守の証明等。
(k) 欧州指令((EU)2016/2102 and(EU)2019/882)に従ったアクセシビリティ要件に適合するようにすること。
(l) EU域外プロバイダー:EU域内代理人を設置すること。

(4) ハイリスクAIシステムに関する輸入者の義務(Article 23)

(a) ハイリスク AI システムを市場に投入する前に、以下を認証することにより当該システムが本規則に適合していること確認しなければならない。
①プロバイダーが Aricle43 に従った適合性評価手続を実施していること。
②プロバイダーが Article11 及び Annex Ⅳに従って技術文書を作成していること。
③CE マークが付され EU 適合性宣言及び使用説明書が添付されていること。
④プロバイダーが権限ある担当者を指名していること。
(b) ハイリスク AI システムが本規則に適合していない又はそれ自体や関連文書に改ざんがあると考える場合、当該システムを市場に出してはならない。
(c) 輸入者の名称、登録商号又は登録商標及び連絡可能な住所を表示しなければならない。
(d) ハイリスク AI システムが輸入者の責任下にある間は、当該システムが本規則に適合しなくなることのないよう保管、輸送しなければならない。
(e) 書類保存義務、情報提供義務等

(5) ハイリスク AI システムに関する販売業者の義務(Article 24)

(a) ハイリスク AI システムを市販する前に、以下を確認しなければならない。
① CE マークが付されていること。
② 必要文書及び使用説明書が添付されていること。
③ プロバイダー及び輸入者が本規則に定める義務を遵守していること。
(b) ハイリスク AI システムが本規則が定める要件を満たしていないと考える場合、当該システムを市販してはならない。
(c) ハイリスク AI システムが販売者の責任下にある間は、当該システムが本規則に適合しなくなることのないよう保管、輸送しなければならない。
(d) 市販したハイリスク AI システムが本規則が定める要件を満たしていないと考える場合、必要な是正措置を講じ、市場から引き揚げ、若しくはリコールし、又は、プロバイダー、輸入者若しくは関係事業者が当該是正措置を講じることを確保しなければならない。
(e) ハイリスク AI システムが、Article 79(1)に規定するリスクをもたらす可能性があると考える理由がある場合、直ちにプロバイダー又は販売業者及び監督当局に通知してそのリスクの内容等を伝えなければならない。

※上記の Article 79(1)に規定するリスクとは、ある製品が、一般的な人の健康と安全、職場における健康と安全、消費者の保護、環境、公共の安全、及び適用される EU 整合法令によって保護されるその他の公共の利益に対し、当該製品の意図された目的に照らし又は当該製品の通常の使用条件又は合理的に予測可能な使用条件(使用期間及び使用開始、設置及びメンテナンスの要件(もしあれば)を含む)のもとで合理的かつ許容可能と考えられる程度を超える悪影響を及ぼす可能性を有していることを意味するものとされています。

(6) ハイリスク AI システムに関する利用者の義務(Article 26)

(a) 使用説明書に従って使用するよう適切な技術的、組織的な措置を講じること。
(b) 必要な能力を有し訓練を受け権限を有する自然人に人的監視措置を委ねること。
(c) インプットデータとハイリスク AI システムの本来の用途との関連性を確保すること。
(d) ハイリスク AI システムの動作を監視し、必要に応じプロバイダーに通知すること。ハイリスク AI システムの使用要領に従った利用が Article 79(1)に規定するリスクをもたらす可能性があると考える理由がある場合、プロバイダー又は販売業者及び監督当局に通知し、システムの使用を停止すること。重大な事故を発見した場合には、直ちにプロバイダーに通知し、その後販売業者又は輸入者及び監督当局に通知する。
(e) ログがコントロール可能な場合、自動的に生成するログを、意図されている使用目的に照らして適切な期間(ただし法に別段の定めのない限り最低6カ月)保存する。
(f) 職場でハイリスク AI システムを使用する前に、利用者は、労働者代表及びその影響を受ける従業員にハイリスク AI システムの対象となることを通知しなければならない。
(g) 公的機関がハイリスク AI システムを利用する場合は、Article 49 に定める登録義務を遵守する。
(h) EU 法により求められるデータ保護インパクト評価を行うことに関する遵守事項

(i) 犯罪被疑者に対する事後的遠隔生体認証のためにハイリスク AI システムを利用する場合の義務
(j) 自然人に関する決定のために AnnexⅢに規定するハイリスク AI システムを利用する場合、当該自然人にハイリスク AI システムの適用を受けることを知らせる。
(k) 本規則の目的を達するために管轄の監督機関が行う措置に協力すること。

(7) 基本的権利影響評価(Article 27)
一定のハイリスク AI システムを利用する一定の利用者について、以下の基本的権利影響評価を実施する義務が課されています。

(a) Annex III のハイリスク AI システム(重要インフラ分野を除く)の利用者(クレジットスコア又は生命保険・医療保険に関するリスク評価に用いられるハイリスク AI システムの利用者である場合を除き、公共機関か公共サービスを行う私企業に限る。)は、利用開始前に、当該システムの利用がもたらす基本権への影響を評価しなければならない。この評価には、以下の要素を含めなければならない。
①当該ハイリスク AI システムを本来の目的のために利用するにあたっての手順
②当該ハイリスク AI システムが利用される期間及び頻度
③当該ハイリスク AI システムの利用によって影響を受ける可能性のある自然人又はその集団のカテゴリー
④当該ハイリスク AI システムの利用により影響を受ける自然人又はその集団が蒙る可能性のある不利益の内容
⑤人的監視措置の概要
⑥上記のリスクが顕在化した場合に講じる措置
(b) 影響評価は、ハイリスク AI システムの最初の使用に適用される。使用中に上記基準に該当する事項が変動し又は最新でなくなったと考える場合、該当の情報を更新するための必要な措置を講じなければならない。
(c) 影響評価の過程で概説されたリスクを軽減するための詳細な計画が特定できない場合、利用者は、ハイリスク AI システムの使用を差し控え、プロバイダー及び当局に遅滞なく通知しなければならない。

## 4. 特定の種類の AI システムに関するプロバイダーと利用者の義務 (Chapter Ⅳ)

AI Act は、上記までのカテゴリーとは別に、特定の AI システムを取り上げて、Chapter IV において以下に紹介する透明性確保等の義務を課しています。

(1) 自然人と直接やりとりする AI システム(Article 50, paragraph 1)
プロバイダーは、当該自然人に AI システムとやりとりしていることを知らせなければならない。
(2) 音声、画像、動画又は文字情報を生成する AI システム(生成 AI システム)(Article 50, paragraph 2)
プロバイダーは、当該生成物が人工的に生成されたものであることが電子的に解読可能な形で当該生成物に表示されるようにしなければならない。
(3) 感情認識システム又は生体分類システム(Article 50, paragraph 3)
利用者は、当該システムが適用される自然人に当該システムが適用されていることを知らせなければならない。
(4) ディープフェイク(AI が生成した画像、音声、動画で、実在のものに似ているため真正なものとの誤信を生ぜしめるもの。Article 3(60))の画像、音声又は動画を生成する生成 AI システム(Article 50, paragraph 4)
利用者は、当該コンテンツが人工的に作成されたものであることを開示しなければならない。
(5) 公衆に公共の利益に関する事項を伝える目的で公表される文字情報を生成する生成 AI システム(Article 50, paragraph 4)
利用者は、当該文字情報が人工的に作成されたものであることを開示しなければならない。
ただし、かかるシステムの利用が、犯罪の探知、予防、捜査又は訴追のために法により認められる場合及び当該生成物が人によるレビュー又は編集を受け、自然人又は法人が当該生成物の公表にあたり編集責任を負う場合には、利用者にかかる開示義務は適用されない。

## 5. 汎用 AI モデル (Chapter Ⅴ)

AI Act は、禁止される AI システム、ハイリスク AI システムとは別に、汎用 AI モデル及びシステミックリスクを有する汎用 AI モデルのカテゴリーを設け、それぞれの AI モデルについてのプロバイダーの義務を規定しています。汎用 AI モデルは、禁止される AI システム、ハイリスク AI システムと並立するカテゴリーではなく、禁止される AI システム、ハイリスク AI システムにも組み込まれ得るものです。

(1) 汎用 AI モデルの定義(Article 3(63))
AI モデル(大規模な自己監視機能を使用して大量のデータで訓練されるものを含む)のうち、有意な汎用性を示し、当該モデルが市場に投入される方法にかかわりなく広範で明確なタスクを適切に実行することができ、様々な下流のシステム又はアプリケーションに統合することができるもの。ただし、市場に投入される前に行われる研究、開発又は試作品製造のために用いられる AI モデルを除く。

(2) システミックリスクを有する汎用 AI モデル
(a) システミックリスクの定義(Article 3(65))
ハイ・インパクトの性能(最も高性能の汎用 AI が有する性能に匹敵又はそれを超える性能)を有する汎用 AI モデルについてのリスクであって、その影響の及ぶ範囲の広さ又は公衆衛生、安全、治安、基本的権利、社会全体に対する現実の若しくは合理的に予見可能な悪影響のゆえに連合市場に重大な影響を及ぼし、バリューチェーン全体にわたって大規模に伝播する可能性のあるもの。
(b) システミックリスクを有する汎用 AI モデルと認定される要件(Article 51)
ア 一定の指標及びベンチマークを含む適切な技術的手法に基づきハイ・インパクトの性能と認められる場合。学習に用いられる計算能力が FLOPs(floating point operations)で計測された結果一定の基準を超える場合、ハイ・インパクトの性能と推定される。
イ 欧州委員会が、一定の手続きにより、AnnexXIII の基準に基づきアと同等の性能を有すると認定した場合。

(3) 汎用 AI モデルのプロバイダーの義務 (Article 53 para1)
(a) AI モデルの技術的文書を作成し最新に保つこと。
(b) (性能及び限界の十分な理解及び本規則の遵守を可能たらしめ、かつ、Annex XII に定める要素を最低限含む)情報及び文書を作成し最新に保ち、汎用目的 AI モデルを AI システムに統合しようとする利用者に対し公開する。
(c) EU の copyright law を遵守するポリシーを制定する。
(d) 汎用目的 AI モデルのトレーニングに用いられたコンテンツに関する十分に詳細なサマリーを作成し一般に公開する。

(4) システミックリスクを有する汎用 AI モデルのプロバイダーの義務(Article 55 para 1)
(a) システミックリスクを特定し軽減するための敵対性試験の実施及び文書化を含めた先端技術を反映した標準化されたプロトコル及びツールにより、モデルの評価を実施する。
(b) EU レベルでの起こり得るシステミックリスクを発生源を含め評価し軽減する。
(c) 重大なインシデント及び取り得る是正措置に関連する情報を遅滞なく追跡、文書化及び報告する。
(d) システミックリスクを有する汎用 AI モデル及びその物理的インフラについて、十分なサイバーセキュリティの保護措置を確保する。

## 6. その他の規定事項 (Chapter VI - Chapter XI)

AI Act は、上記 III. 1.(Chapter I)から 5. (Chapter V)のとおり AI システムの類型毎に、各当事者の義務、遵守事項を定めるほか、イノベーション支援の方法(Chapter VI)、ガバナンス(Chapter VII)、ハイリスク AI システムのための EU データベース(Chapter VIII)、市場投入後のモニタリング、情報共有、市場監視(Chapter IX)、行動規範及びガイドライン(Chapter X)、権限の委任及び委員会手続(Chapter XI)等について定めています。

## 7. 罰則 (Chapter XII)

(1) 加盟国は、AI Act の定める条件に従って事業者(プロバイダー、製品製造者、利用者、認定代理店、輸入

業者、販売業者)による AI Act 違反に対する罰則及びその他の執行措置の詳細を定めるものとされています(Article 99)。

(2) かかる罰則の基準は以下のとおり AI Act において定められています。

(a) 禁止行為違反(Article 99 第 3 項): Article 5 に規定される AI システムに係る禁止に違反した場合、最高 3,500 万ユーロ又は違反者が企業の場合は前事業年度の全世界における年間売上高の 7%のいずれか高い方の制裁金を課すものとする。

(b) AI Act のその他の規定違反(Article 99 第 4 項): AI システムが AI Act の以下の規定(Article 5 に規定されるものを除く)に違反している場合、最高 1,500 万ユーロ又は違反者が企業の場合は前事業年度の全世界における年間売上高の 3%のいずれか高い方の制裁金の対象とする。

① Article 16 に基づくプロバイダーの義務。

② Article 22、Article 23、Article 24 又は Article 26 に基づく、認定代理店、輸入業者、販売業者又は利用者の義務。

③ Article 31、Article33(1)、Article33(3)、Article 33(4) 若しくは Article34 に基づく通知先機関の要件及び義務。

④ Article52 に基づくプロバイダーと利用者の透明性義務。

(c) 不正確、不完全又は誤解を招く情報の提供(Article 99 第 5 項):要請への応答として不正確、不完全、又は誤解を招くような情報を通知先機関若しくは国家の管轄当局に提供した場合、750 万ユーロ以下の罰金又は違反者が企業の場合は前事業年度の全世界における年間売上高の 1%のいずれか高い方の制裁金の対象とする。

(d) 中小企業に対する閾値の引き下げ(Article 99 第 6 項):中小企業(スタートアップを含む。)の場合、上記の制裁金は、記載されたパーセンテージ又は金額のいずれか低い方を上限とする。

(e) 汎用 AI モデルのプロバイダーに対する制裁金(Article 101):欧州委員会は、汎用 AI モデルのプロバイダーが故意又は過失により以下に該当した場合、前事業年度の全世界における年間売上高の 3%若しくは 1,500 万ユーロのいずれか高い方を上限とする制裁金を課すことができる:

① AI Act の適用規定に違反した場合。

② Article 91 に基づく文書又は情報の要請に従わず、又は不正確、不完全若しくは誤解を招くような情報を提供した場合。

③ Article 93 に基づき要請された措置を遵守しない場合。

④ 欧州委員会に対して所定の汎用 AI モデルへのアクセスを付与しない場合。

## 8. 今後の手続き

冒頭に記載したとおり、AI Act は欧州議会の承認を得た上で本年 5 月 21 日に EU 理事会においても承認されたため、今後欧州連合官報に掲載され、かかる掲載の 20 日後に発効し、(特定の条項を除き)その 2 年後から適用が開始されます[3]。

以上

弁護士 後藤 出 オブ・カウンセル
izuru.goto@city-yuwa.com

弁護士 池辺 健太 パートナー
kenta.ikebe@city-yuwa.com

シティユーワ法律事務所

〒100-0005 東京都千代田区丸の内 2-2-2 丸の内三井ビル 7 階

[3] https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/