2023 年 11 月 30 日　全 8 頁

バイデン米大統領、AI の安全性に係る大統領令に署名

連邦政府機関に AI の開発や利用に係る基準等の策定を求める

金融調査部 研究員 矢田歌菜絵

[要約]

- 2023 年 10 月 30 日にバイデン米大統領は、AI の安全性に係る大統領令「安全かつ安心で信頼性のある AI の開発と利用に係る大統領令」（以下、AI 大統領令）に署名した。米国ではこれまでに、個別の政府機関による基準が公表されたり、州法レベルでの AI 規制が施行されたりしているが、AI 大統領令は連邦レベルでの対応の集大成である。

- AI 大統領令は、主に政府機関に基準の策定等の対応を求めるもので、次の実践のための重点項目に分けられる。(1) AI の安全性とセキュリティに係る新しい基準、(2) 米国民のプライバシーの保護、(3) 公平性と市民権の推進、(4) 消費者・患者・学生の保護、(5) 労働者の支援、(6) イノベーションと競争の促進、(7) 諸外国における米国のリーダーシップの強化、(8) 連邦政府による効率的かつ責任ある AI の利用の保証。

- 民間企業においては、AI の開発に際し、政府機関に安全性テストの報告等が必要になる。ただし、EU の AI 規制（案）のように一部の AI の開発・利用を禁止するものではなく、AI 大統領令は政府機関が策定する基準やベストプラクティスに基づいた対応を求めるにとどまる。

- 米国議会では AI 規制に係る法整備が進められており、2023 年 11 月 15 日に上院で超党派の法案が提出された。これは、AI 大統領令からさらに踏み込んだ内容であり、法案の成立も含め、今後の動向が注目される。

大統領令の公表

2023 年 10 月 30 日にバイデン米大統領は、AI（人工知能）の安全性に係る大統領令「安全かつ安心で信頼性のある AI の開発と利用に係る大統領令」(Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence)（筆者仮訳、以下、該当同じ）[1]（以後、本稿では AI 大統領令とする）に署名した。

[1] The White House "Executive Order on the Safe, Secure, and Trustworthy Development and Use of

米国においては、AI の開発および利用に関して、連邦レベルでの包括的な法規制はないものの、連邦政府機関は様々な対応を進めてきた。2022 年にバイデン政権が「AI 権利章典の青写真」(Blueprint for an AI Bill of Rights) [2]を発表し、AI 開発等の際に米国民を保護するための「5 つの原則」[3] (法的拘束力なし) を示した。その後、連邦取引委員会 (FTC) [4]や米著作権局 [5]等の個別の連邦政府機関が AI に対する基準等を発表している。一方、州レベルでの対応について見ると、少なくとも 15 の州においては AI 規制に係る法律が 2023 年 9 月末時点で施行されており、その他の州でも法案が提出されている。米国において、こうした背景がある中、今般、連邦レベルでの対応の集大成として、大統領令が署名されるに至った。

AI 大統領令の概要

ポリシーと原則

バイデン政権では、図表 1 で示した指針や優先事項に基づいた AI の開発および利用の促進・管理を政策として掲げている。

図表 1 指針・優先事項

(a) AIが安全かつ堅牢でなければならない。
(b) 責任あるイノベーションや競争、協同を促すことで、AI分野での米国のリードを後押しし、困難な社会課題の解決可能性を高める。
(c) 責任あるAIの開発および利用は、米国の労働者の支援に徹する。
(d) AI政策は、バイデン政権の公平性や市民権に対する取り組みと整合的でなければならない。
(e) AIやAI対応サービスを利用・購入する米国民の利害は守られなければならない。
(f) AIの発展につれ、米国民のプライバシーおよび自由権は守られなければならない。
(g) 米国民にとって有益となるよう、連邦政府において、AI利用から生じるリスクの管理や、責任あるAI利用の規制、管理、および支援する内部の能力を拡大することは重要である。
(h) 米国が過去にイノベーションでリードしてきたことを踏まえ、連邦政府はグローバルにおける社会的かつ経済的、技術的な進歩を先導する必要がある。

(出所) 前掲脚注 1　AI 大統領令セクション 2 「政策および原則」より大和総研作成

Artificial Intelligence" (2023 年 10 月 30 日)
https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

[2] The White House "Blueprint for an AI Bill of Rights: A Vision for Protecting Our Civil Rights in the Algorithmic Age" (2022 年 10 月 4 日)
https://www.whitehouse.gov/ostp/news-updates/2022/10/04/blueprint-for-an-ai-bill-of-rightsa-vision-for-protecting-our-civil-rights-in-the-algorithmic-age/

[3] この 5 つの原則とは、「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データのプライバシー、④ユーザーへの通知と説明、⑤人による代替手段、配慮、フォールバック」のことである。
(出所) 内閣府「人間中心の AI 社会原則会議 (令和 4 年度 第 2 回)」(2022 年 12 月 21 日) 資料 3「米国の AI 権利章典 (AI Bill of Rights) について (内閣府提出資料)」
https://www8.cao.go.jp/cstp/ai/ningen/r4_2kai/2kai.html

[4] Federal Trade Commission "Keep your AI claims in check" (2023 年 2 月 27 日)
https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check

[5] U.S. Copyright Office "Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence" (2023 年 3 月 16 日)
https://copyright.gov/ai/ai_policy_guidance.pdf

実践のための重点項目

AI 大統領令と同時に公表されたファクトシート[6]では、AI 大統領令を実践のための重点項目に分けて説明している（図表 2）。いずれも連邦当局に対して基準やベストプラクティスの策定、一定の対象者に報告義務を課す等の緩やかな規制であり、EU の AI 規則（案）[7]のように非常にリスクの高い AI を禁止する等の AI の利用そのものを制限するものではない。

図表 2 AI 大統領令を実践するための重点項目

(1) AIの安全性とセキュリティに係る新しい基準
・ AIシステムの開発者に、連邦政府への安全性テストの結果等の共有を義務付け
・ 国家の安全保障に重大なリスクを及ぼす可能性のある基盤モデルを開発する企業に、連邦政府への公開前の安全性テストの結果の共有を義務付け
・ AIで生成されたコンテンツへの認証や「電子透かし」に関するガイドラインを制定
(2) 米国民のプライバシーの保護
・ プライバシー保護技術の開発および利用への連邦政府による支援を強化
(3) 公平性と市民権の推進
・ AIの不適切な利用による差別や偏見等を防ぐため、社会保障プログラムや政府請負業者、家主等に対して、AIアルゴリズムによる差別を助長させないような明確な基準を策定
(4) 消費者、患者、学生の保護
・ 医療分野：AIを適切に使った、安価で命を救う薬剤の開発を促進
・ 教育分野：AIツールを活用する教育者を支援
(5) 労働者の支援
・ AIによる不当な監視や偏見、解雇等のリスクを低減させ、メリットを最大化させるための原則とベストプラクティスを策定
(6) イノベーションと競争の促進
・ 医療や気候変動に関するAI研究には助成金を拡大
・ 中小の開発者・企業にも機会のある公正かつ開かれた、競争的なAIエコシステムを促進
(7) 諸外国における米国のリーダーシップの強化
・ 諸外国のパートナーと重要なAI基準の開発・適用を加速させ、相互運用性のあるものに
(8) 連邦政府による効率的かつ責任あるAIの利用の保証
・ AI利用のガイドラインの制定

（出所）ファクトシート（前掲脚注 6）より大和総研作成

図表 2 で示した実践のための重点項目を基に、AI 大統領令の主要な規定の概略は次の（1）から（8）の通りである。

（1）AI の安全性とセキュリティに係る新しい基準

AI 大統領令では、産業界のコンセンサス形成を促すため、商務省や国立標準技術研究所（NIST）に、安全かつ安心で信頼できる AI システムの開発導入に係る基準やベストプラクティスを定めることを要求している。この基準やベストプラクティスの作成にあたっては、生成 AI に関する

[6] The White House "FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence"（2023 年 10 月 30 日）
https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

[7] European Commission "Regulatory framework proposal on artificial intelligence"（2023 年 11 月 15 日更新）
https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

既存の AI リスクマネジメントのフレームワーク（NIST AI 100-1）[8,9]の補完的な内容も含めている。加えて、特にサイバーセキュリティやバイオセキュリティの分野において、AI の性能を評価・監督するためのガイダンスや指標を策定するためのイニシアチブを取ることも求めている。また、AI 開発者が AI の安全性テスト（レッドチームオペレーション[10]）を適切に行えるよう、適切な基準の策定も求めている。

特に原子力、生物、化学、重大なインフラ、エネルギー分野において、悪影響を及ぼし得る結果を生成する可能性のある AI については、AI セキュリティリスクの低減を目的として、その性能を評価するためのツールの作成が関係当局に求められている。

AI の安全性および信頼性を確保するために、国防生産法に基づき関係当局は、国家の安全保障に重大なリスクをもたらすおそれのある基盤モデルの開発企業に、そのモデルの開発やトレーニングに係る計画や状況の報告を求めることになる。具体的には、モデルウエイトや高度な脅威に対するトレーニングプロセスの物理的保護措置やサイバーセキュリティの保護措置に関するものも含まれる。また、NIST が策定する基準に基づいて実施される安全性テストの結果の報告も対象となる企業に求められる。

大規模コンピューター・クラスターを取得、開発または保有する企業や個人等に対しては、そのコンピューター・クラスターの保有や場所、処理能力等に関する報告の関係当局への提出が求められる。

米国における IaaS（Infrastructure as a Service）プロバイダーは、外国人がそのプロバイダーのサービスを通じて悪意あるサイバー活動のおそれのある大規模な AI モデルのトレーニングを行う場合、その外国人の身元やそのようなサイバー活動の存在等について関係当局への報告が求められる。

AI によって生成されたコンテンツによるリスクを低減するため、連邦政府またはその代理が生成されたコンテンツであることの識別や「電子透かし」等によるラベル付けを促進する、と記載されている。そのために当局は、既存のツールや実例に係る報告書を作成し、基準を策定し、適切かつ整合的な範囲で当該基準の現行規制への盛り込みも検討するとある。具体的には、コンテンツの認証や生成元の確認、「電子透かし」等の生成されたコンテンツであることのラベル付け、生成されたコンテンツであるとの識別、生成 AI によって子供の性的虐待になるようなコンテンツや当人の同意なく実在人物の画像（身体の一部の精細なデジタル表現を含む）の生成の防止、そしてこれを行うためのソフトウェアのテスト、生成されたコンテンツの監督管理につ

[8] National Institute of Standards and Technology "NIST Trustworthy and Responsible AI - NIST AI 100-1"（2023 年 1 月 26 日）
https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

[9] このフレームワークは、AI の設計や開発、展開、利用を行う組織にとって、AI に係る多くのリスクを管理し、信頼でき、かつ責任のある AI の開発および利用を促進するためにリソースを提供することを目的としている。AI に係るリスクの考え方や信頼できる AI の特徴、AI に係るリスク管理等について説明している。

[10] 様々な攻撃を仕掛けることで、セキュリティ対策の実効性を確かめるもの。詳細は、大和総研「レッドチームオペレーション -WOR(L)D ワード」（https://www.dir.co.jp/world/entry/red-team-operation）を参照されたい。

いても検討される。

(2) 米国民のプライバシーの保護

AI によるプライバシー（個人に関する情報の収集や利用、推論含む）へのリスク低減のため、連邦機関が作成した商用利用可能なデータ（特に個人が特定できるデータ）を特定・評価し、そのデータの収集・加工・維持・利用・共有・配布・処理に関する当局の基準や過程を検証し、当局の基準の修正を検討する、と記載されている。

(3) 公平性と市民権の推進

AI による不当な差別や他の実害を防ぐため、関連機関と連携して現行法のエンフォースメントを強化する、と記載されている。現行法に従い、適切な範囲で、捜査官や検察官に対して自動化システム、AI を含む市民権侵害および差別に関する調査および起訴のベストプラクティスに関するガイダンス、技術支援、および学習の機会の提供が検討される。

特に不動産市場において、自動ツールやアルゴリズムによって生じる不当な差別や偏見をなくすため、追加的な基準の策定が定められている。特に犯罪歴や立退き歴、信用情報等のデータの利用は連邦法に抵触する可能性があり、それを防ぐためのベストプラクティスについても検討する、としている。

(4) 消費者、患者、学生の保護

AI の利用による詐欺や差別、プライバシーへの脅威から米国民を保護するため、AI の利用から生じ得る他のリスク（金融安定性リスクを含める）に対処するため、適切と認められる場合、独立規制機関による全権の行使が推奨されている。また、ルール制定のみならず、既存規制の枠組みに AI が適用できないか、規制当局による明確化も推奨されている。

医療分野では、安全かつ責任ある AI の導入および利用の確保のために、関係当局はタスクフォースを設置し、ポリシーや枠組みを作成する。また、薬剤開発での AI の利用を規制する戦略を立てる、としている。

教育分野においては、AI に係るリソース・ポリシーや基準が作成される。また、教育指導者のための AI ツールキットも作成する（信頼性・安全性の向上や教育に係るプライバシーに関する法令規則との整合性を目的に、人間による AI の判断に対するレビューや、教育に特化したガードレール[11]の作成も含む）としている。

[11] AI のリスクを抑え、適切に利用するための基準を指す。

(5) 労働者の支援

大統領経済諮問委員会の委員長はAIの労働市場への影響に係る報告書を大統領に提出し、労働省は、AIの導入によって解雇された人への政府機関による支援の可能性を大統領に報告する必要がある。

また、AIを導入した職場における従業員のウェルビーイングを確保するために、関係当局は、従業員のウェルビーイングに対するAIの潜在的悪影響を低減させ、潜在的利益の最大化を目的とした、雇い主に参照されるような原則とベストプラクティスを作成および公表する。ベストプラクティス策定後、関係機関は、現行の法律に則って、適切な範囲でガイドラインの採用を推奨することが規定されている。

(6) イノベーションと競争の促進

AI関連労働者や研究者のビザの発行促進や、国家AI研究リソース(the National AI Research Resource: NAIRR)のパイロットプログラム等を定めている。

また、医療分野や気候変動分野における研究に重きを置くことについても記載している。

競争の観点からは、政策規制当局は、現行法での適切な範囲において、AIの競争を促進することができるようになる。特に、FTCにおいては、AI市場における公正な競争を確保し、AIの使用によって生じる可能性のある損害から消費者と労働者を保護するために、適切と判断された場合には既存の権限の行使を検討することが推奨されている。半導体産業における競争やイノベーションも促進するとしている。AIイノベーションやAI商用化に取組む中小企業に対しては、支援機関を設立し、資金を優先的に配分するとしている。

(7) 諸外国における米国のリーダーシップの強化

軍事・情報分野以外での同盟国や国際的パートナーとの取組み拡大に努める、と記載されている。二国間や、多国間の利害関係者のフォーラムにおいて、米国の既存および計画中のAIに関する基準と政策に関する理解を深め、国際的な協力を促進する、とある。

AIに関する基準(AIに関する命名法や用語、データの収集・処理・保護・プライバシー・機密性・取扱い・分析、AIシステムの信頼性・検証・保証、AIリスクマネジメント)の開発および推進の国際的な取組み計画を策定するとしている。

(8) 連邦政府による効率的かつ責任あるAIの利用

全ての連邦機関は、最高AI責任者(Chief AI Officer)を指名し、連邦機関間のAI評議会が設置される。最高AI責任者は、所属機関における最高の責任を負い、他の連邦機関最高AI責任者と協力し、AIの利用を調整、所属機関でのAIイノベーションの促進、当該機関におけるAI利用のリスクを管理する等の責任を負うことになる。

連邦政府における安全かつ責任ある生成 AI の利用を推進するため、連邦機関は必要に応じてリスクマネジメントに基づいて特定の AI サービスの利用を制限し、適切な生成 AI の利用のためのガイドラインと制限を設けることになる。また、適切なリスク対策措置を講じたうえで、少なくとも米国民の権利にほとんど影響を及ぼさないと考えられる試験的業務や日常的業務に関しては、安全かつ信頼できる生成 AI へのアクセスを職員やプログラムに許可してもよいとされている。

連邦政府の情報を保護するため、連邦機関はリスクマネジメント（職員に連邦政府の情報の適切な利用、保護、伝達、破棄に関するトレーニングを受けさせる等）を行うとしている。

今後の動向

米国

AI 大統領令は EU の AI 規則（案）のように特定の AI の開発や利用を禁止するのではなく、主に関係当局に対して調査、大統領への報告、基準・ベストプラクティスの策定を求めるものであるといえる。目下、当局への報告義務（前述「実践のための重点項目（1)」）が課される可能性が高い民間企業はごく一部と言われている。

連邦政府によって生成されたコンテンツに対する「電子透かし」等の導入（前述「実践のための重点項目（1)」）は 2024 年の大統領選において、有害なディープフェイク等による混乱を防ぐ目的があるとも言われている。

AI 大統領令は主に連邦政府機関に対応を定めるものであり、民間企業への影響は政府機関を通した間接的なものである。その内容も政府機関が策定した基準に則った対応を民間企業に求めるものであり、EU のように一部の AI の利用を禁止する等の厳格な対応を求める規制とはいえないだろう。ただし、これはあくまでも AI 大統領令における規定であり、今後米国において AI 規制に関する法律が整う可能性がある。

2023 年 11 月 15 日に、AI の開発や利用に係る法案[12]が超党派で上院に提出された。この法案は AI 大統領令の一部の内容について、より踏み込んだものとなっている。

この法案では、重要なインフラ等の最もリスクが高いとされる AI の利用（the highest-risk application of AI）に対して保護措置を講ずることを盛り込んでいる。具体的には、重要なインフラや生態認証といった最もリスクが高いとされる AI を開発する企業に対して、その企業が特定した AI システムのリスクやその軽減・制御策をサービス展開前に当局に報告する必要があると記載されている。

また、同法案では、AI が生成したコンテンツに対する「電子透かし」等のラベル付けだけで

[12] Library of Congress "S.3312 - Artificial Intelligence Research, Innovation, and Accountability Act of 2023"（2023 年 11 月 15 日）
https://www.congress.gov/bill/118th-congress/senate-bill/3312

なく、人間が作成したコンテンツに対してもクレジットを付与するための開発についても言及している。特に、大手インターネットプラットフォームに対しては、AI で生成されたコンテンツである旨を利用者に明確で分かりやすく提供するよう求めている。

同法案の成立も含め、今後の米国での法整備が注目される。

日本

国際的には、2023 年 5 月に開催された G7 広島サミットを受けて、2023 年 10 月 30 日に「広島 AI プロセスに関する G7 首脳声明」[13]が発表された。これは、同時に発表された AI 開発者向けの国際的なルール作りの枠組みである「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指針」および「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」を G7 として歓迎するものである。今後、「広島 AI プロセス包括的政策枠組」の年内の策定に向けて作業が進められている。

日本国内においては、2023 年 10 月 31 日に経済産業省「AI 事業者ガイドライン検討会」の第 1 回会合 [14]が開催され、「AI 事業者ガイドライン」の策定に向けた議論が進められてきている。2023 年 11 月 7 日に開催された内閣府「AI 戦略会議」の第 6 回会合 [15]では、「AI 事業者ガイドライン」の原案が会合内で示されている。「AI 事業者ガイドライン」の原案は非公表であるため、「AI 戦略会議」の公表資料 [16]から推測すると、「AI 事業者ガイドライン」も AI を開発および利用する事業者に対して一定の AI の利用を禁止する等の厳格な対応を求めるもの、というよりもむしろ、一定の基準を示し、その基準に則ることを促す対応が主たるものとして見受けられる。

日本においても、AI の開発や利用における基準の策定が進められており、AI 開発者や利用者に対して今後より厳格な対応が求められるか注目される。

---

[13] 外務省「広島 AI プロセスに関する G7 首脳声明」(2023 年 10 月 30 日)
https://www.mofa.go.jp/mofaj/ecm/ec/page5_000483.html
[14] 経済産業省「第 1 回 AI 事業者ガイドライン検討会」(2023 年 10 月 31 日)
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/2023_001.html
[15] 内閣府「AI 戦略会議　第 6 回」(2023 年 11 月 7 日)
https://www8.cao.go.jp/cstp/ai/ai_senryaku/6kai/6kai.html
[16] 前掲脚注 15 資料 1-3「AI 事業者ガイドライン等の行動規範の履行確保及び AI 利用の促進の検討について(案)」