AIに起因する選挙リスクとAIガバナンス 米国調査レポート
AI Election Risks and US AI Governance Measures

2024年12月23日

情報処理推進機構

IPA Better Life with IT

目次

調査内容は2024年10月31日時点のものです。Trump次期大統領に関わる事態進展は含まれません

1．報告アウトライン

- 調査の経緯：AIのセキュリティ脅威と対策検討には、AI利用で先行する米国の実態調査が不可欠

  2021年度　初回調査（深層学習ブーム、脅威実態）

  2023年度　第2回調査（生成AI以降の脅威実態）

  ⇒　2024年度5月に公開

  2024年度　第3回調査（選挙への脅威、政府のガバナンス対策分析）

  重要インフラ事例として選挙に注目

- 調査手法

  米国調査会社への委託

  文献調査

  ガバナンス・AI・セキュリティに関する有識者インタビュー

調査報告書目次1

調査報告書目次2

調査報告書目次3

Executive summary

調査の概要

1　米国国内の選挙に関わるAI脅威の実態調査

　　重要インフラとしての選挙に対する脅威の実態

2　米国のAIガバナンス政策の有効性分析

　　バイデン大統領令、AI権利章典，AIリスク管理フレームワーク、他ガイドラインの有効性分析（独自に提案する評価手法に基づく）

2．調査結果要旨

2.1 脅威の実態（4章）（2023年度調査結果再掲）

脅威	Risk リスク	影響をうける主体 etc.	時期 1]	影響
AI-enhanced traditional cyberattacks	Force multiplier for disruptive attacks	All sectors but critical infrastructure may be impacted greatly	Medium-term	High
AIで強化された従来のサイバー攻撃	Increased capabilities, sophistication, and efficiency of cybercriminals in ransomware and cryptocurrency-related cyberattacks; lowered barrier to entry	Individuals and industries, especially ransomware-prone industries such as health care, financial, and hospitality sectors	Medium term	High
	Li フィッシングの容易化・高速化 social e ency and speed in spear phishing	Individuals, industries, governments, academia, news organizations, critical infrastructure	Immediate	High
AI-enabled disinformation	Domestic Disinformation: increased ce 国内の虚偽情報言動監視 ulnerable gr arian digital norms	Particularly individuals and minorities in authoritarian nations, democracy, freedom of speech	Immediate	Medium
AIを利用した虚偽情報	Sta 国家支援虚偽情報キャンペーン ation car societies, erosion of trust in institutions, degrading of democracy	Individuals, democratic governments, electoral process Democratic	Immediate	Medium
	Promotion of crime and discrimination: new class of crime such as deepfake pornography and stock market manipulation	Individuals, finance industry, black market, private sector widely	Medium term	Medium
	Ele 選挙妨害 ıe censorship, disinformation	Individuals, freedom of speech, democratic nations, electoral process	Immediate	Medium-High

AI-Enabled disruption or maloperation of s AIによるシステム障害とAIシステムへの攻撃	Data poisoning: false outputs leading to bac データポイズニング nation, dis	Critical infrastructure, social infrastructure, justice system, others	Medium term	High
	Inherent biases and vulnerabilities: rei データバイアスによる意思決定妨害 content ger ng	Individuals, businesses, governments	Immediate	Medium
	Inte failures: ope 不正出力による誤判断・誤作動 se out	Critical infrastructure, social infrastructure, justice system, multiple industries	Immediate	Medium-High
AI-enabled national security threats AIによる国家安全保障上のリスク	Military applications: potential aut 自動兵器・戦闘意思決定支援 , military dec nical concerns	Defense sector, governments	Long term	High
	AI i テスト不十分な軍用導入 ith unp n	Governments, defense sector, industry	Long term	High
	Espionage and Mass Surveillance: hig 諜報・大規模監視 eous uses by the private sector	Public and private sector, individuals, privacy	Medium term	Medium
	Ter テロリズムプロパガンダ pro ans	Social media companies, individuals, governments	Medium term	Low
	Bioterrorism: development of novel pathogens, efficient information gathering	Individuals, healthcare, and pharmaceutical sectors	Long term	Low
Business risks due to misuse of 生成AI誤用によるビジネスリスク	Vul nd dis 脆弱なコード流通 非倫理的・不正コンテンツ生成・レピュテーションリスク ins ing): data lea , reg ıancial losses, operational disruption	Businesses, consumers, employees, privacy	Immediate	Medium
	Lega data leak 法的リスク 営業秘密漏えい non lties	Legal system, privacy, businesses, individuals	Immediate	Medium

2.1 脅威の実態（4章）（2023年度調査結果再掲）

- セキュリティリスク認知
  - AIは技術・利用とも急激に変化しており、リスクの全体像はまだ見えていない
  - 利用者側のリスクの自覚と対応は米国でも今一歩
- 悪用の傾向：
  - 「既存の攻撃を早く、大量に」から始まっている
- 脅威の大きさ：
  - フェイクコンテンツはすぐわかる低品質のものでも、大量にばらまかれると厄介
  - マルウェア生成を含む攻撃自動化は時間がかかるが、急速にできる可能性も
- 脅威の対応：
  - 大規模学習データが正しい（汚染されていない）ことの検証は難課題
  - 利用者の誤用・悪用を抑止する枠組みが必要

2.2 選挙の脅威（5章）

- 内容のサマリ

大統領選挙に関しては、生成AIによる新しい脅威は生まれていない

一方で、世論分断等のサイバー脅威はすでに存在し、生成AIはそれを増幅している

海外からの選挙干渉は増加したが、AIによる大規模な影響は観測されていない

フェイク蔓延によるデジタル空間不信の脅威（Liar’s dividend: うそつきの報酬）は存在し、ガザ戦争では顕在化している

技術に依存しない多面的な対策が必要である

2.2 選挙の脅威（5章）

- 対応策

  既存対策の強化が基本的な考え方

  - Proactive preparedness

    選挙関係コンテンツはオーサライズしたものを流通させる

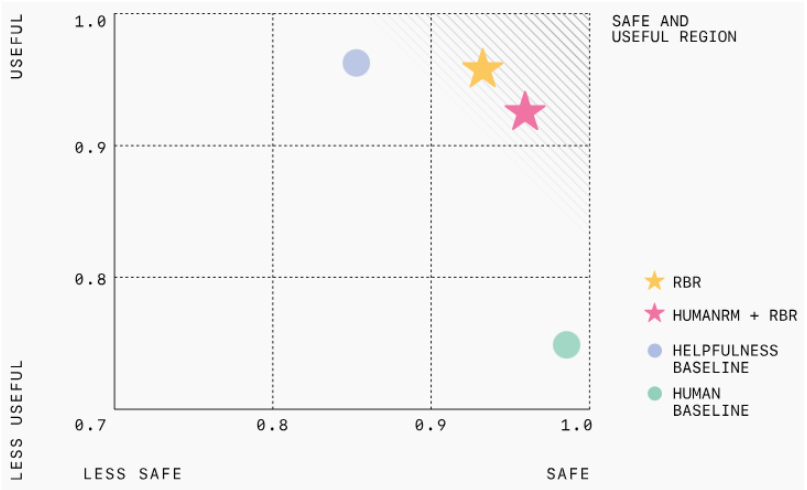
    メディアとのパートナーシップ
  - Policy and government

    既存法制の効果的活用（金融分野等の手法展開）
  - Cybersecurity Tools and Technical Methods

    フィッシング対策強化（政府機関のメールセキュリティ等）
  - AI hardening

    生成AI自身のセキュリティ強化、レッドチーミング

    フェイク検知等のAI活用



報告書 Figure4: Ensuring Safety and Utility in OpenAI's Technologies

2.3 主要なガバナンス対策（6章）

- 内容サマリ
  - 主要対策として　AI 権利章典（AI-BoR）、大統領令（EO14110）、リスクマネジメントフレームワーク（AI-RMF）、NISTのAI関連規格（SP800-218A等）を精査
  - リスクベースのAI RMF と 人権を意識したEO14110が柱
  - これらが2023年度調査で抽出した５つの脅威をカバーしていることを確認。しかし、
    - 各ガイドラインのカバーはパッチワーク的で整理されてはいない
    - 中心となるガバナンスの枠組みはなく、強制力がはたらきにくい
    - 全ての脅威について、ケースの具体化や対策は示されていない

2.3 ガバナンス対策の脅威別カバレッジ（6章）

- 緑（十分カバー）、黄（一部カバー）、グレー（カバーされない） の3レベル

Threat / Governance Approach	EO 14110	AI BoR	NIST AI RMF	NIST SSDF for GenAI (SP800-218A)
AI-Enhanced Traditional Cyberattacks AI強化従来攻撃	Well covered—emphasizes cybersecurity, including measures for offensive cyber operations and guidelines for auditing AI capabilities to mitigate potential harm.	Adequately covered—aims for safe and secure systems which would mitigate against traditional cyberattacks.	Not covered.	Well covered—Focuses on protecting software from unauthorized access and producing well-secured software with minimized vulnerabilities.
AI-Enabled Disinformation & Misinformation 虚偽情報	Well covered—focus on safe and secure systems as well as defending against related cyberattacks can defend against false information, requires DOC to work on authenticating GenAI content.	Adequately covered—suggests human alternatives and fallbacks which can filter false information.	Well covered—Confabulation, toxicity and bias, and obscene content risks.	Adequately covered—by ensuring a safe and secure AI development lifecycle, hallucinations can be mitigated which will lead to less misinformation.
AI-Enabled Disruption or Maloperation of Systems 破壊攻撃	Well covered—requires agencies to coordinate to develop guidelines that ensure critical infrastructure systems' resiliency suggests red teaming and other assessments to ensure quality AI systems, and more.	Adequately covered—algorithmic discrimination protection, human alternatives and fallbacks, and safe and effective systems can help mitigate system disruptions.	Adequately covered—identifies data integrity, information security, and dangerous recommendation risks.	Adequately covered—by ensuring a safe and secure AI development lifecycle, disruption and maloperation of systems can be mitigated.

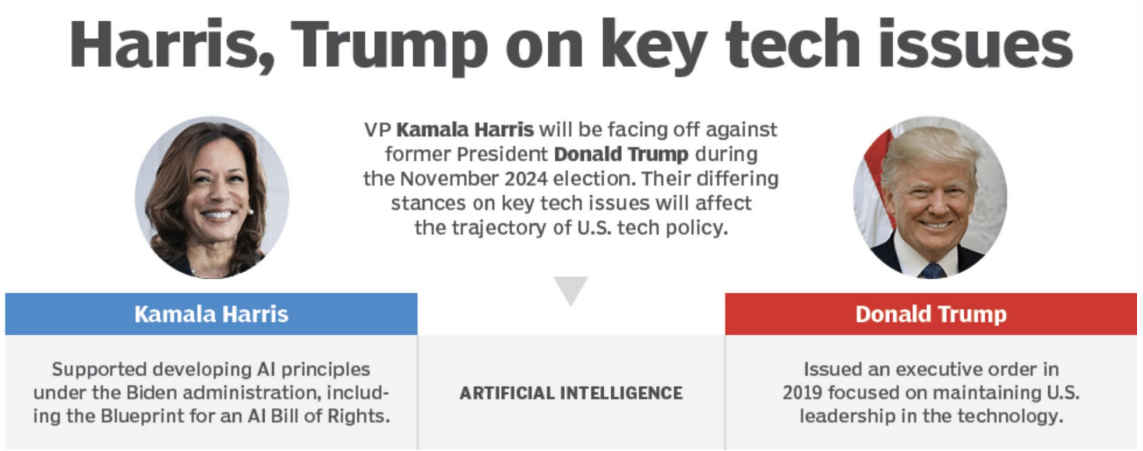
2.3 ガバナンス対策の脅威別カバレッジ（6章）

AI-Enabled National Security Threats ナショナルセキュリティ	Well covered—AI EO is the US's foundational strategy for AI which is largely aimed at protecting the nation from CBRN risks to critical infrastructure risks to more.	Adequately covered—safe and secure systems can mitigate national security risks, and civil rights/liberties and privacy can help mitigate against espionage.	Adequately covered—identifies CBRN weapons risk.	Adequately covered—by ensuring a safe and secure AI development life cycle, eventually impacts national security.
Business Risks Due to Misuse of GenAI 生成AI誤用	Well covered—much of the DOC's work is related to securing businesses from potential GenAI risks.	Well covered—allowing users access to equal opportunities and resources, protecting user data and privacy, mitigating against algorithmic discrimination, and requiring safe and secure systems can help reduce business risks. (機会均等・アルゴリズム差別の禁止を明示)	Well covered—identifies data privacy, intellectual property, dangerous recommendations, value chain integration, and Human-AI interaction risks. (プライバシー、知財、危険な推奨、インタラクション等のリスクを明示)	Well covered—refers to risks in the AI development lifecycle from preparing the organization to responding to vulnerabilities. (AI開発ライフサイクル全般にわたる脆弱性対応を明示)

誤用は各対策で重点ポイントあり、
総合的にカバー

参考：EO14110のインパクト

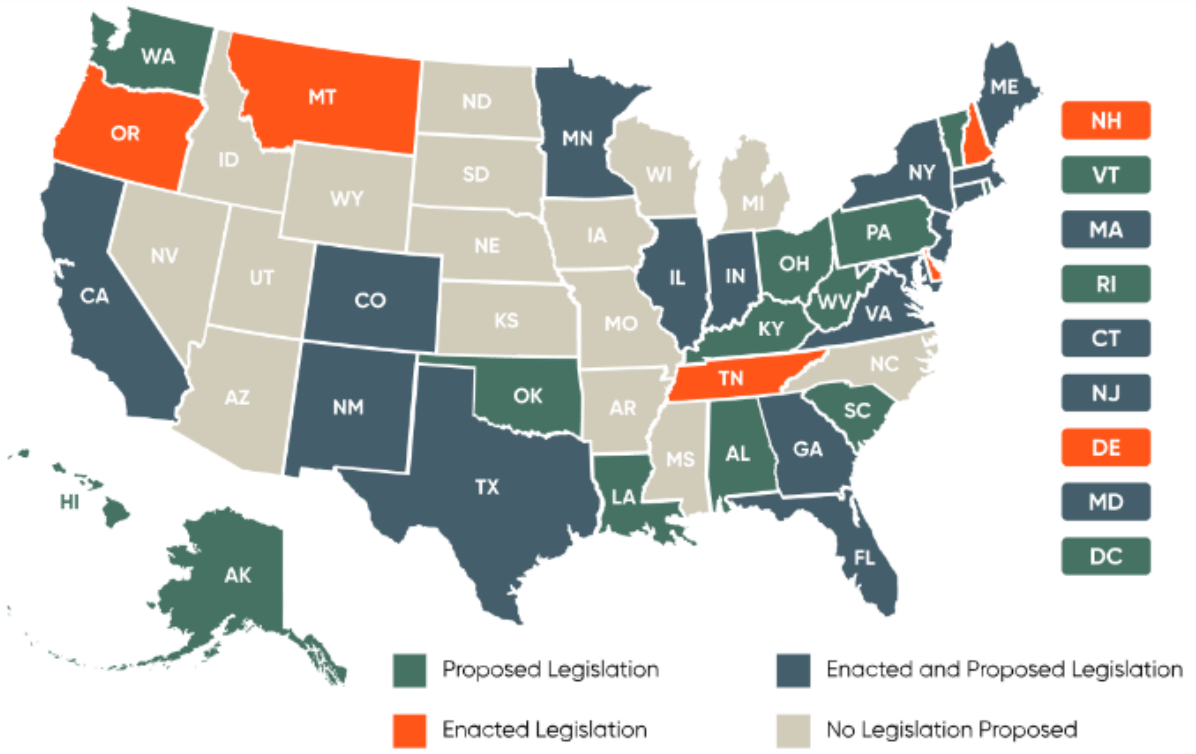
- EO14110の特徴
  - AIと民主主義原則との紐づけ
  - リスクベース（AI RMF）と人権ベースのハイブリッドアプローチ
  - 有害性除去にフォーカス
  - 民間の参画
- トランプ次期大統領のAI方針は？
  - 重視は変わらないが、言論の自由を盾にテストを制限する可能性



報告書 Figure 9: Comparison of 2024 Presidential Candidates' Views on AI

参考：各州の規制（6章）

- カリフォルニア州の状況
  - 2024年9月、Newsom州知事は大規模AIモデル開発者に安全テスト・被害発生時責任を課す法案（SB1047）に拒否権を行使
  - 大規模AIの効果的な規制の枠組みはステークホルダ間で継続検討へ
- 各州共通の課題
  - 透明性・説明性の強化
  - 消費者保護
  - 雇用差別
  - 倫理のプラクティス
  - 規制とビジネスのバランス



報告書 Figure 10: Overview of State-by-State AI legislation

2.4 対策評価とまとめ：指標の導出（7章）

- 米国の対策・インタビュー・評価の枠組事例をもとに、3レベルの評価指標を導出
  - Pillar1 基本原則（分野横断）
    - マルチステークホルダの関与
      - 社会、産業界、分野エキスパート)
    - 説明性
    - 適応性　定期的な技術・リスクの見直しと修正
      - サイバーセキュリティが先例を示している
  - Pillar2 高レベル戦略（分野横断）
    - 冗長性　　各種リスクへの対処
    - 国際的整合性
    - 包括的リスク管理　AIライフサイクル全般にわたるマネジメント（人間中心、社会的責任の重視）
  - Pillar3 ユースケース（分野別の規制）
    - 分野フォーカス　　分野固有の規則導入（金融、ヘルスケア、交通運輸・・）
    - 利用・成果ベース　具体的成果を目標とする
    - 長期的な整備　　ケース・規則は時間をかけて作成（サイバーセキュリティが先例）

報告書 Figure 17: Pillars to Develop Effective AI Governance

2.4 対策評価とまとめ：評価例（7章）

- NIST SP800-218Aを前頁指標で評価した例

Pillar 1: Foundational Principles	Pillar 2: Higher-Level Strategies	Pillar 3: Sectoral Use Cases
Multistakeholder Engagement: NIST typically engages various stakeholders for its publications, and the EO required NIST to solicit private sector, academia, and public sector input.	**Robustness**: covers various risks at the software development level, aimed at software producers and acquirers only.	**Sectoral Focus:** none.
Explainability: aligned with business mission requirements, organizational goals, risk tolerance, and available resources; helps identify gaps and guide a prioritized action plan.	**International Harmony:** refers to ISO documents.	**Use and Outcome-based:** only defines risks at a high-level, not use case focused.
Adaptability: based on software development practices (BSA, OWASP, SAFECode)	**Comprehensive Risk Management:** focuses on software security but spans the AI model development process (data sourcing to training to software integration); does not include AI deployment or operation	**Long-term Development:** NIST tends to publish follow ups and updates to their documents.

緑　充足
黄　一部充足
赤　未充足

分野別ケースが課題であることが可視化された

2.4 対策評価とまとめ：脅威と対策の対応（8章）

- 2023年ど調査で5分類した脅威とガバナンス対策の対応

脅威	AIガバナンス対策（ガイドライン）	その他の規格・対策	課題・制限
AIで強化された従来攻撃（フィッシング等）	EO14110	既存のサイバーセキュリティ規格・対策	攻撃増加
AIで強化された虚偽情報	選挙関連法制、各州法 AI BoR	生成AIコンテンツ検知・ラベリング・トラッキング	規制がDeepFakeに偏る 多層防御 嘘つきの報酬
AIで強化された破壊的攻撃	EO14110 AI BoR	既存のサイバーセキュリティ規格・対策	ケース策定・テスト
ナショナルセキュリティ脅威	AI RMF AI600-1	AI100-5他の国際連携規格	ケース策定・テスト
生成AI誤用	AI BoR AI RMF SP800-218A		倫理テスト 利用リテラシー ベンダー説明

2.5 有識者インタビュー要旨（9章）

- インタビュー対象者
  - AI・IT政策専門家（1）
  - セキュリティコンサルタント
  - AI・IT政策専門家（2）
  - 情報・民主政策専門家
  - AI政策研究者
  - AI・AI政策研究者
  - 政府技術政策ディレクター
  - 選挙アナリスト
  - AI・AI政策研究者
  - データ政策コンサルタント

インタビュー内容：AI・IT政策専門家（1）

- 大統領令やAI RMFでは、責任あるAI開発とプライバシー保護を重視しているが、包括的AI政策の欠如と規制調整の複雑さが課題である。
- 包括的なAIアプローチには、国際連携推進が不可欠であり、民間のニーズを理解し、自主規制のによるイノベーションを国際慣行と一致させることが重要。
- AIガバナンスは急速な技術進歩に対応するため、能力開発と透明性の向上、柔軟な規制メカニズムの導入が必要である。
- 政府・企業はサンドボックス環境で規制と製品開発を共同でテストする。シンガポールや英国などが実装で先行している。
- NISTは、シンガポールや英国と協力してAIのリスク管理フレームワークや、公平性、説明責任、透明性、信頼性、セキュリティなどの基準を設定している。
- 効果的なAIガバナンスには、定期的な監査、プライバシー評価、継続的なバイアステストが必要であり、透明性と説明責任、利害関係者間のパートナーシップの確立が不可欠である。

インタビュー内容：セキュリティコンサルタント

- AI脅威はまだ明確ではない（サイバーセキュリティ初期に類似）。セキュリティとAIの相違は成熟度と知識不足。NISTサイバーセキュリティフレームワーク等で、人々はセキュリティスク対応を学んだ。AIも同じ道をたどるだろう。
- ガバナンスフレームワークはアカデミックな傾向があるので注意。AIによる標準ビジネス、は通用しない（セキュリティでは失敗）。機能しなければ柔軟に変更すべき。
- 今は原則があって体験がない。よい利用・だめな利用のケースが重要。ケースが足りない、と規制にとびついてはいけない。よいケースを想定し、過剰な規制をさけるべき。スタンフォード大学では「問題がないAI画像」のDB実装を行っている。よいケースの指標が必要だろう。
- 知識のない利用者はケースに依存しがちだが、それが正しい保証はない。よいケース、悪いケースの混在は10年スパンで続く。悪いケースを排除せず、バランスをとるべき。
- 選挙セキュリティについては①ベースレイヤ（全体投票集計等）、②フェイクニュース拡散、③実投票の結果集積、の3レイヤがあるが、②③は対応が遅れている。一方で、州別のインフラになるので攻撃が連鎖するリスクは小さい。

インタビュー内容：AI・IT政策専門家（2）

- AI規制の最大課題は技術の不確定性。リスクは製品安全性を超え、人権が含められて複雑化した。継続的なモニタリングが必須である。
- 生成AIがリスクを過大に意識させ、規制圧力がかかった。不必要な規制は避けるべき。
- リスクアセスメントは変化の早い技術ではなく、ケースで行うべき。マルチステークホルダプロセスが有効で、欧州は実践を始めている。ハイリスク・ローリスク双方のケース共有が有効。ハイリスクだけではネガティブな影響が出る。
- 生成AIでは適切な学習・誤用への備えが必要。各ステークホルダに責任があるが、製造者の責任は学習の説明、ガイダンス等で大きい。実践は規制や業界ルールなどの選択肢がある。
- AIの影響のアセスメントにはあらゆるセクターの参加が重要。官民連携では、統制された環境でサービス試行を行うサンドボックスがありうる。
- 欧州では企業・団体・消費者の専門家グループが活動し、早く、信頼できる規制につながると考える。参加する専門家について、今はサプライサイドのバリューチェーンのプレーヤが主体である。

インタビュー内容：情報・民主政策専門家

- 選挙脅威は2016年のロシアの直接的干渉から、2020年の投票ツールに関する不正疑惑など、ボトムアップにナラティブを拡散させ、既存の対立を拡大する戦略に変化した。
- 台湾総選挙では選挙に関するフェイク画像が確認されたが効果は小さく、欧州も同様だった。米国でもバイデン大統領のフェイク音声が確認されたが影響は小さかった。また既存のナラティブ強化にもAIは用いられたが、大きな影響はなかった。
- 対策としては、情報空間に信頼できる情報を提供し続けることだと言える。そのためにも多層防御が重要である。信頼できるサイトへの誘導、ラベリング・透かしによる信用付与、AI合成検知、投票者のリテラシー教育・責任分担が必要である。
- メディアリテラシーに対して米国は十分投資していない。資金だけではなく、教育リソースや政府政策が重要で、時間をかけて行うべき。
- 嘘つきの報酬は深刻な問題である。重要な対策は信頼できる情報をたどれること（content provenance）。情報ソースへの信頼は民主主義の根幹である。
- 海外の敵対勢力は国内の論争・対立を助長させる戦略をとるため、国内・海外の脅威の区別はしづらくなっている。

インタビュー内容：AI政策研究者

- AI政策に関する大統領令の実装を過去2件の大統領令を含め調査しているが、EO14110の実装は省庁の優先度が高く、うまく進んでいる。ただし実装の情報公開はまだ十分でない。
- ここまでの実装は体制やChief AI Officerアサイン等の基盤作りで、今後の施策にとって重要である。またOMBによる連邦政府のAI利用ポリシー、NIST/CISAの規格作成等は特筆される。
- ＥＯ14110は現行機関が適所でAI強化を行うスキームであるため、組織間連携は明示的でない。Chief AI Officerのボード等が今後想定される。
- 大統領令の執行にはAI人材が必須であるが、連邦政府内のAI人材の確保・強化は課題である。移民政策、外部技術支援、民間連携などの施策が必要となる。
- 大統領令はまだボランタリーな枠組であり、今後各省庁が民間拘束力のある規制をどう作るか、が課題である。

インタビュー内容：AI・AI政策研究者（１）

- AI RMF、EO14110は「試行錯誤」アプローチをとり、整理されてはいない。EO14110はボランティアベースで放任に見えるが、CISAが重要インフラ向けガイドラインを出しており、実装は進むと思われる。
- AI RMFも任意だが、アップデートがなされ、そのタイミングも見える点が長所である。一方、ケースや実績が見えていない点が課題である。さらに実装の情報がなく、どの機関が提供するかも不透明で、このギャップを埋めることが必要。
- NISTはユースケース作成、日本政府とのAI RMFクロスウォーク、US AISI立ち上げ・官民連携コンソーシアムによるレッドチーミング、ベンチマーキングを行っている。
- US AISIの活動では、民間参加者への情報提供が局所的でリスクの全体像がわからない、等の情報共有に課題がある。
- 他国のAISIにおいても、政府から民間への情報提供が課題解決に重要である。官民連携のフィールドとしては技術支援モデルや開発プロセスも有用である。

インタビュー内容：政府技術政策ディレクター

- 金融分野のAIリスクはバイアス・公平性が重要でガイダンス策定を急いでいる。AIセキュリティはフィッシング等とフェイク情報を重大リスクとみる。サプライチェーンリスクも懸念点である。多くのサービスにAIが入り込むことを想定しないといけない。
- 財務省の金融向けAIサイバーリスクマネジメントガイドはよくできている。
- AI RMFは対策が概論的だがこれでよい。実装にステークホルダが参画すればよい。技術進歩が早いので止めるべきでない。実装の主体となる政府機関はCISA、民間はOWASPか。
- 組織のガバナンスは、集権的なChief AI Officerより委員会形式がいい。連携の形式はサイバーセキュリティの知見が役にたつ。
- ユースケースはSMB（中小企業）がからむサービスが重要である。サイバーセキュリティではSMB対応が問題化し、AIでもこれを懸念する。
- AI監査、経営陣のリスク把握、経営陣とのコミュニケーションが非常に重要である。
- AIが心理に及ぼす影響のテストは難しい。フェイクや公平性についてここからだめ、と境界をひくのは難しく、「より安全な選択肢」を選ぶ方法をつくるしかないだろう。

インタビュー内容：選挙アナリスト

- 生成AIとボット・マイクロターゲティング等が組み合わされ、悪意の虚偽情報の拡散スピード・規模が一変した。
- 米国において政治団体等の生成AI利用はめだたない。民間の利用の影響は精査中。VoIP SNSでネイティブ・ラテン系アメリカ人への攻撃が行われ、証拠が残りにくい問題がある。
- 各州で虚偽情報規制法制ができているが、成立は難しい。法案はDeepfakeに偏っている。連邦政府は生成AI脅威の演習手法を公開した。LLMによる悪意の応答は、高齢者等に脅威である。規制は必要で、単発でなく多段の政策が重要である。
- 生成コンテンツ保証技術としてウォーターマークと来歴標準はカリフォルニア州で法制化されたが、実装は難しい。虚偽情報生成は国内外どちらもあり、海外は中国が主である。EUの選挙ではEU域内の虚偽情報が主、との報告がある。
- 嘘つきの配当（Liar's dividend）と呼ぶ状況（情報がみな虚偽に見える）はガザ戦争で起きている。選挙には致命的で、デジタルリテラシー向上プログラムが必須である。
- 選挙脅威は虚偽情報のほか、選挙管理部門の誤用、投票検証用AIのセキュリティがある。フィッシングも課題である。
- リスク軽減策として、包括的プライバシー保護法を望む声がある。生成AIコンテンツ検知は有用だが、効果は不透明である。ファクトチェックも有用だが、ナラティブを信じた人への効果には課題がある。

インタビュー内容：AI・AI政策研究者（２）

- ニューヨーク市の包括的AI戦略はデータインフラ、AIアプリ、AIガバナンス/ポリシー、パートナーシップ、ビジネスと経済開発を包含し、テーマ領域として、AI教育コンテンツ提供、市内AIエコシステムのマッピングを目的としている。
- 効果的なAIガバナンス対策はドメインやアプリケーションで異なるため、画一的なアプローチは避けるべきで、ガバナンスフレームワークは柔軟であるべき。
- NISTとCISAによるAI安全対策はボランタリーで、高リスクなシステムの対応が遅れている。AIの悪用を抑制するのに効果的とはいえない。政府機関の連携・調整を強化し、フレームワークの施行を促進すべきである。
- 市の組織は多大な経済損失・レピュテーション被害を受け、サイバーセキュリティ規範の必要性を認識している。AIガバナンスでは、その複雑さを感じながらも厳格な措置と暗号化プロトコルの採用等を進めている。
- AIガバナンスは、ドメイン固有のアプローチがある。主観を伴う社会政治的問題や管理方法に関するトレードオフが存在し、環境（コンテキスト）やアプリケーションへの依存性もあるため、より複雑である。

インタビュー内容：データ政策コンサルタント

- AI RMFの特徴は柔軟性・冗長性、既存フレームワークの活用、マルチステークホルダアプローチである。弱点は技術志向・詳細になりうる点で、政策議論には向かない。
- EO14110はよくできているが濃淡がある。プライバシー等は既存規則に頼っている。最も重要な関連活動はプラクティス・SSDFによるフレームワークの具体化である。
- AI RMFはNISTサイバーセキュリティフレームワークと同様セクターで共有すべき。問題は利用者の知識がないこと。ここはセキュリティと異なる。
- AIガバナンスフレームワークで必要なのはリスクベース、環境依存性。汎用モデルではリスクがいえない。リスクアセスメントでは、学習・開発・利用のサプライチェーンリスクが課題。もう１つのリスクは誤用。技術の説明性は非常に重要。
- 政府機関のAIリスク規制については、現行の技術中立な法制が機能している（金融、医療等）。ただし機関を横断するケースで問題がおこりうる（FTCとCFPBなど）。
- AISIの課題は規則の境界定義である。多くのAI応用が始まると境界は複雑になる。
- Chief AI OfficerはAIリスクの統括者で、高次の統合リスク管理組織に位置づけられるのがよい（CISOとは異なる）。

2.5 有識者インタビューサマリ1

- ガバナンスフレームワーク（AI RMF, EO14110）の全般評価
  ボランティアベースであり、出発点としてはよい
  概念的だがそれでよい
  技術進歩が速いので拙速な規制をすべきでない
  領域別ケース・知識不足が課題で、現状では具体的施策が弱い

2.5 有識者インタビューサマリ2

- フレームワーク具体化の施策
  - サイバーセキュリティ対策の経験が役に立つ
  - 利用者のAI知識向上策が必要
  - どの機関が具体化を担うか、の可視化は重要
  - 具体化はマルチステークホルダアプローチが有効
  - 実際は「きれいな標準」ではすまないので、柔軟な変更が必要
  - 官民連携では必要な情報共有が重要
  - 生成AIの誤用ではベンダーの利用者への説明が重要

2.5 有識者インタビューサマリ3

- 選挙セキュリティに関するAI脅威
  生成AIによる大統領選挙への直接脅威は顕在化していない
  「嘘つきの配当」の不安は大きく、対策が必要
  LLM悪用（VoIP等の誘導）、虚偽情報検知は決定的な対策がない
  検知・追跡技術だけでなく、リテラシー向上を含む多層的な対策が必要

Thank you!