AISI
Japan AI Safety Institute

AI セーフティ ファクトシート 2025

Factsheet of AI Safety in Japan 2025

2026 年 3 月 31 日

AI セーフティ・インスティテュート（AISI）事務局

目次

1　主要なAI戦略、アクションプラン、制度

本ファクトシートは、AI セーフティを検討するにあたり公開された文書や設立された新たな組織および体制を中心に国内外の事実関係を記載したものである。

１．１　原則、戦略

1) 人間中心の AI 社会原則[1]（2019 年/統合イノベーション戦略推進会議）

AI 戦略実行会議の下、AI をより良い形で社会実装し共有するための基本原則を検討し、AI 戦略に反映させることを目的として設置され、原則が策定された。このような考察の下で、特定の技術やシステムが「AI」かを区別するのではなく、広く「高度に複雑な情報システム一般[2]」がこのような特徴と課題を内包すると捉え、社会に与える影響を議論した上で、AI 社会原則の一つの在り方を提示し、AI の研究開発や社会実装において考慮すべき問題が列挙されている。

2) 統合イノベーション戦略（2024 年、2025 年/内閣府）

統合イノベーション戦略 2024[3]には、「先端科学技術の戦略的な推進」、「知の基盤（研究力）と人材育成の強化」、「イノベーション・エコシステムの形成」の３つの基軸で政策を推進していくとともに、３つの強化方策として、「重要技術に関する統合的な戦略」、「グローバルな視点での連携強化」、「AI 分野の競争力強化と安全・安心の確保」が推進される。

また、統合イノベーション戦略 2025[4]には、２．（１）先端科学技術の戦略的な推進①重要分野の戦略的な推進 として、AI イノベーション促進とリスク対応の両立、AI の研究開発の推進、AI 関連施設等の整備及び共用の促進、AI 活用の推進、AI 適正性の確保、AI 関連人材の確保と教育振興、AI に関する調査研究、AI 分野の国際的協調の推進が記載されている。

---

[1] https://www8.cao.go.jp/cstp/ai/aigensoku.pdf
[2] この宣言の中では、AI のことを「高度に複雑な情報システム一般」と同一視していく。
[3] https://www8.cao.go.jp/cstp/tougosenryaku/togo2024_zentai.pdf
[4] https://www8.cao.go.jp/cstp/tougosenryaku/togo2025_zentai.pdf

１．２　原則や戦略に関する国際合意や主要レポート

1)OECD AI principles（2019 年/OECD）[5]

OECD AI 原則は、AI に関する初の政府間基準である。人権と民主的価値を尊重する革新的で信頼性の高い AI の推進に貢献してきている。2019 年 5 月に採択され、後述の広島 AI プロセスを踏まえ、2024 年 5 月に改訂された。この原則は、政策立案者と AI 関係者に実用的かつ柔軟な指針を提供する 5 つの価値観に基づく原則と 5 つの加盟国政府等が政策等に反映させるべき勧告で構成されている。

2)UNESCO Ethics of Artificial Intelligence（2021 年/UNESCO）[6]

ユネスコで 2021 年 11 月、AI 倫理に関する世界初のグローバルスタンダードとなる「人工知能の倫理に関する勧告」が採択された。人権と尊厳の保護は勧告の要であり、透明性や公平性といった基本原則の推進を基盤とし、常に AI システムの人間による監視の重要性を念頭に置いている。

3)広島 AI プロセス（2023 年/G7）[7]

2023 年５月に開催された G7 広島サミットの結果を踏まえ、その急速な発展と普及が国際社会全体の重要な課題となっている生成 AI について議論するために、2023 年５月に立ち上げられた。

全ての AI 関係者向けの 広島プロセス国際指針（2023 年/G7）[8]

2023 年 10 月に公表された「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」を基礎として、安全、安心で信頼できる AI の実現に向けて、AI ライフサイクル全体の関係者それぞれが異なる責任を持つという認識の下、同指針の 11 項目を、高度な AI システムの設計、開発、導入、提供及び利用に関わる全ての関係者に適宜適用し得るものとして整理した上で、偽情報の拡散等の AI 固有リスクに関するデジタルリテラシーの向上や脆弱性の検

---

[5] https://www.oecd.org/en/topics/ai-principles.html
[6] https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
[7] https://www.soumu.go.jp/hiroshimaaiprocess/
[8] https://www8.cao.go.jp/cstp/ai/ai_senryaku/7kai/11hiroshimaaipurosesu.pdf

知への協力と情報共有等、利用者に関わる内容が 12 番目の項目として追加されている。

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（2023年/G7）[9]

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は、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に基づき、安全、安心で信頼できる AI を世界に普及させることを目的とし、最先端の基盤モデル及び生成AIシステムを含む、最も高度なAIシステムを開発する組織による行動のための自主的な手引きとして提供される。

国際行動規範の「報告枠組み」に係る合意（2024年/G7）[10]

「国際行動規範」を自主的に遵守する AI 開発企業等による履行状況を確認するための手法の基本的な運用方法及び質問票に関して、2024 年 12 月に合意された。「国際行動規範」に沿って作成されたAI開発企業等への質問票をOECDの Web サイト上で公開し、企業が回答する。「報告枠組み」の運用開始。なお、2026年3月現在、日本企業9社を含む25組織が回答を提出し、OECDのウェブサイト上で公表されている。

4)Final Report of the High-level Advisory Body on Artificial Intelligence - Governing AI for Humanity（2024年/国際連合）[11]

国連事務総長直属の「AI に関するハイレベル諮問委員会」の最終報告書「人類のためのAIガバナンス」が2024年9月に発表された。報告書では、AIの国際ガバナンスの促進に関する提言が記載されている。

5)UN Global Digital Compact（2024年/国際連合）[12]

2024年9月22日、「未来サミット」の成果文書である「未来のための約束」の附属文書として、すべての人々にとって包摂的かつオープンで持続可能、公平、安全で安心なデジタルの未来を推進するための目的、原則及び行動を定めるグローバル・デジタル・コンパクトが採択された。

---

[9] https://www.soumu.go.jp/hiroshimaaiprocess/documents.html
[10] https://www.soumu.go.jp/menu_news/s-news/01tsushin06_02000306.html
https://transparency.oecd.ai/reports
[11] https://www.un.org/en/ai-advisory-body
[12] https://www.un.org/techenvoy/global-digital-compact

１．３　法令

日本は、ガイドラインを中心としたソフトローでAIセーフティが推進されてきた。一方で、法令の必要性、法律を策定する場合の対象や実施事項なども検討されている。

1)AI関連法

a)人工知能関連技術の研究開発及び活用の推進に関する法律（AI法）[13]

生成AIをはじめとするAI技術の発展は、国民生活の向上及び国民経済の発展に寄与する一方、国内の AI 開発・活用は遅れており、また、多くの国民がAIにより発生するリスクに不安を抱えている状況である。

AIのイノベーションを促進しつつ、リスクに対応するため、令和７年６月４日にAI法が公布・一部施行され、９月１日にはAI戦略本部の設置に係る規定等も含め、全面施行された。

2)デジタル関連法

a)デジタル社会形成基本法[14]

デジタル社会の形成が、我が国の国際競争力の強化及び国民の利便性の向上に資するとともに、急速な少子高齢化の進展への対応その他の我が国が直面する課題を解決する上で極めて重要であることに鑑み、デジタル社会の形成に関する施策を迅速かつ重点的に推進し、もって我が国経済の持続的かつ健全な発展と国民の幸福な生活の実現に寄与するため、デジタル社会の形成に関し、基本理念及び施策の策定に係る基本方針、国、地方公共団体及び事業者の責務、デジタル庁の設置並びに重点計画の作成について定められた。

b)デジタル原則に照らした規制の一括見直しプラン[15]

我が国がデジタル化を図っていく上での指針となるべき「構造改革のた

---

[13] https://www8.cao.go.jp/cstp/ai/ai_act/ai_act.html
[14] https://laws.e-gov.go.jp/law/503AC0000000035
[15] https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_re

めのデジタル原則」を策定し、当該原則に適合したデジタル社会の実現を目指して、各府省庁とも連携し、構造改革に取り組んでいく計画が示された。

3) 個人情報・知的財産関連法

a) 個人情報保護法等[16]

個人情報の有用性に配慮しながら、個人の権利や利益を守ることを目的とした法律である。この法律は、デジタル社会の進展に伴い個人情報の利用が著しく拡大していることに鑑み、個人情報の適正な取扱いに関し、基本理念及び政府による基本方針の作成その他の個人情報の保護に関する施策の基本となる事項を定め、国及び地方公共団体の責務等を明らかにし、個人情報を取り扱う事業者及び行政機関等についてこれらの特性に応じて遵守すべき義務等を定めるとともに、個人情報保護委員会を設置することにより、行政機関等の事務及び事業の適正かつ円滑な運営を図り、並びに個人情報の適正かつ効果的な活用が新たな産業の創出並びに活力ある経済社会及び豊かな国民生活の実現に資するものであることその他の個人情報の有用性に配慮しつつ、個人の権利利益を保護することを目的とされる。

b) 知的財産に関する法令

特許法[17]：発明の保護及び利用を図ることにより、発明を奨励し、もつて産業の発達に寄与することを目的とされる。

著作権法[18]：著作物並びに実演、レコード、放送及び有線放送に関し著作者の権利及びこれに隣接する権利を定め、これらの文化的所産の公正な利用に留意しつつ、著作者等の権利の保護を図り、もつて文化の発展に寄与することを目的としている。

商標法[19]：商標を保護することにより、商標の使用をする者の業務上の

---

sources/cb5865d2-8031-4595-8930-8761fb6bbe10/e3650360/20220603_meeting_administrative_research_outline_07.pdf

[16] https://www.ppc.go.jp/personalinfo/

[17] https://laws.e-gov.go.jp/law/334AC0000000121

[18] https://laws.e-gov.go.jp/law/345AC0000000048

[19] https://laws.e-gov.go.jp/law/334AC0000000127

信用の維持を図り、もつて産業の発達に寄与し、あわせて需要者の利益を保護することを目的とされる。
意匠法[20]：意匠の保護及び利用を図ることにより、意匠の創作を奨励し、もつて産業の発達に寄与することを目的とされる。
不正競争防止法[21]：事業者間の公正な競争及びこれに関する国際約束の的確な実施を確保するため、不正競争の防止及び不正競争に係る損害賠償に関する措置等を講じ、もって国民経済の健全な発展に寄与することを目的とされる。

4) 自動運転関係の法令

道路交通法[22]：道路における危険を防止し、その他交通の安全と円滑を図り、及び道路の交通に起因する障害の防止に資することを目的とした法令。
道路運送車両法[23]：道路運送車両に関し、所有権についての公証等を行い、並びに安全性の確保及び公害の防止その他の環境の保全並びに整備についての技術の向上を図り、併せて自動車の整備事業の健全な発達に資することにより、公共の福祉を増進することを目的とされる。

5) ドローン関係の法令

航空法[24]：航空機の航行の安全，航空機による運送事業などの秩序の確立を目的とする法令。航空機の航行の安全及び航空機の航行に起因する障害の防止を図るための方法を定め、航空機を運航して営む事業の適正かつ合理的な運営を確保して輸送の安全を確保するとともにその利用者の利便の増進を図り、並びに航空の脱炭素化を推進するための措置を講じ、あわせて無人航空機の飛行における遵守事項等を定めてその飛行の安全の確保を図ることにより、航空の発達を図り、もつて公共の福祉を増進することを目的とされる。

航空法の他、小型無人機等禁止法、道路交通法、民法、電波法、都道府県の条例等によって規制される。

---

[20] https://laws.e-gov.go.jp/law/334AC0000000125
[21] https://laws.e-gov.go.jp/law/405AC0000000047
[22] https://laws.e-gov.go.jp/law/335AC0000000105
[23] https://laws.e-gov.go.jp/law/326AC0000000185
[24] https://laws.e-gov.go.jp/law/327AC0000000231

6)特定電気通信役務提供者の損害賠償責任の制限及び発信者情報の開示に関する法律[25]

特定電気通信による情報の流通によって権利の侵害があった場合について、特定電気通信役務提供者の損害賠償責任の制限及び発信者情報の開示を請求する権利について定めるとともに、発信者情報開示命令事件に関する裁判手続に関し必要な事項を定める目的とする法令。

7)サイバーセキュリティ基本法[26]

我が国のサイバーセキュリティに関する施策に関し、基本理念を定め、国及び地方公共団体の責務等を明らかにし、並びにサイバーセキュリティ戦略の策定その他サイバーセキュリティに関する施策の基本となる事項を定めるとともに、サイバーセキュリティ戦略本部を設置すること等により、同法と相まって、サイバーセキュリティに関する施策を総合的かつ効果的に推進し、もって経済社会の活力の向上及び持続的発展並びに国民が安全で安心して暮らせる社会の実現を図るとともに、国際社会の平和及び安全の確保並びに我が国の安全保障に寄与することを目的とされる。

サイバー基本法の他、刑法、不正アクセス行為の禁止等に関する法律等によって規制される。

1．4 指針、ガイドライン

1)AI 事業者ガイドライン（2024 年/総務省、経済産業省）[27]

AI の安全安心な活用が促進されるよう、我が国における AI ガバナンスの統一的な指針が 2024 年 4 月に示されている。これにより、様々な事業活動において AI を活用する者が、国際的な動向及びステークホルダーの懸念を踏まえた AI のリスクを正しく認識し、必要となる対策を AI のライフサイクル全体で自主的に実行できるように後押しし、互いに関係者と連携しながら「共通の

---

[25] https://laws.e-gov.go.jp/law/413AC0000000137
[26] https://laws.e-gov.go.jp/law/426AC1000000104
[27] https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_1.pdf
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_3.pdf

指針」と各主体に重要となる事項及び AI ガバナンスを実践することを通して、イノベーションの促進とライフサイクルにわたるリスクの緩和を両立する枠組みを、積極的に共創していくことを目指している。
2024 年の初版以降、2025 年 3 月に日本語第 1.1 版、2025 年 4 月に英語第 1.1 版が公開されている。

2)AI セーフティに関する評価観点ガイド（2024 年/AISI）[28]

AI システムの開発や提供に携わる者が AI セーフティ評価を実施する際に参照できる基本的な考え方を提示している。本書の作成に際しては、日本において AI を活用する事業者が適切に AI を活用するための指針を示す「AI 事業者ガイドライン」を参考としている。
2024 年の初版以降、2025 年 3 月に日本語第 1.10 版、英語第 1.10 版が公開されている。

3)AI セーフティに関するレッドチーミング手法ガイド（2024 年/AISI）[29]

AI システムの開発や提供に携わる者が対象の AI システムに施したリスクへの対策を、攻撃者（AI システムの悪用や破壊を意図する者）の視点から評価するためのレッドチーミング手法に関する基本的な考慮事項を示している。
2024 年の初版以降、2025 年 3 月に日本語第 1.10 版、英語第 1.10 版が公開されている。

4)AI 時代の知的財産権検討会「中間とりまとめ」[30]**（2024 年/内閣府知的財産戦略推進事務局）**

「AI 時代の知的財産権検討会 中間とりまとめ」（2024 年 5 月）は、生成 AI と知的財産権の関係についてガイドラインとしての役割を果たす資料である。「中間とりまとめ」のポイントを、権利者の視点から紹介、解説した「権利

[28] https://aisi.go.jp/output/output_information/250328_1/
[29] https://aisi.go.jp/output/output_information/250331_1/
[30] https://www.kantei.go.jp/jp/singi/titeki2/chitekizaisan2024/0528_ai.pdf

者のための手引き[31]」が 2024 年 11 月に公表されている。

5)AI と著作権に関する考え方について[32](2024 年 3 月/文化審議会著作権分科会法制度小委員会)

生成 AI と著作権の関係についてクリエイターや AI 開発事業者等の懸念の払拭に向け、文化審議会著作権分科会法制度小委員会において議論し、現行の著作権法における考え方を整理したものである。また、本文書の内容を踏まえ、関係当事者の立場ごとにわかりやすくまとめた「AI と著作権に関するチェックリスト&ガイダンス」が 2024 年 7 月に文化庁より公表されている。

6)知的財産推進計画 2025[33]

2025 年 6 月には、知的財産戦略本部より知的財産推進計画 2025 が発表された。同計画では、AI 技術の進展を踏まえた発明等の保護について、AI 利用発明の発明者の定義等の諸論点について早期に結論を得ることを求めており、人工知能（AI）の時代の知的財産制度について、前進した内容となっている。

7)総務省重点施策 2026 [34]

総務省は、令和 8 年度の重点分野として、総務省が積極的に取り組む施策について「デジタル変革を通じた持続可能な地域社会と強い経済基盤の実現（総務省重点施策 2026)」として取りまとめた。「国内外における AI ガバナンスの実現」に取り組むとしており、AI 事業者ガイドラインの更新・周知と広島 AI プロセスの推進について明記された。

8)人工知能関連技術の研究開発及び活用の適正性確保に関する指針（令和 7

---

31 https://www.kantei.go.jp/jp/singi/titeki2/chitekizaisan2024/2411_tebiki.pdf

32 https://www.bunka.go.jp/seisaku/chosakuken/aiandcopyright.html

33 https://www.cas.go.jp/jp/seisakukaigi/titeki2/index.html

34 https://www.soumu.go.jp/menu_news/s-news/01kanbo05_02000197.html

年 12 月 19 日人工知能戦略本部決定）[35]

同年 12 月に、AI 法第 13 条に基づき、全ての AI に関連する主体における AI の研究開発・活用の適正な実施に係る自主的かつ能動的な取組を促すための「人工知能関連技術の研究開発及び活用の適正性確保に関する指針」（令和 7 年 12 月 19 日人工知能戦略本部決定）が策定された。

１．５　その他の政府機関の AI 関連ガイドライン

1)機械学習品質マネジメントガイドライン（2020 年/産業技術総合研究所）[36]

本ガイドラインは、機械学習を利用したシステム、特にその中に含まれる機械学習で実装されたソフトウェアコンポーネント（機械学習要素）の品質に関する基準と達成目標を定めることにより、企業が自ら構築した AI を利用するシステムの品質を測定し向上させ、また AI の誤判断による事故や経済損失などを減少させる一助となることを目的とされたものである。2020 年の初版以降、2023 年 12 月に日本語第 4 版、2023 年 1 月に英語第 3 版が公開されている。また、2025 年 5 月に「生成 AI 品質マネジメントガイドライン第 1 版」が公開されている。

2)初等中等教育段階における生成 AI の利活用に関するガイドライン[37]（2024 年/文部科学省）

2024 年 12 月に発行された本ガイドラインは、教職員や教育委員会等の学校教育関係者を主たる読み手として、学校現場における生成 AI の適切な利活用を実現するための参考資料となるよう、利活用に当たっての基本的な考え方や押さえるべきポイントをまとめたものである。生成 AI の概要、基本的な考え方を示した上で、学校現場において押さえておくべきポイントとして、利活用する場面や主体に応じた留意点について、現時点の知見が基に可能な限り具体的に示されている。

---

[35] https://www8.cao.go.jp/cstp/ai/ai_guideline/ai_guideline.html

[36] https://www.digiarc.aist.go.jp/publication/aiqm/

[37] https://www.mext.go.jp/content/20241226-mxt_shuukyo02-000030823_001.pdf

3) コンテンツ制作のための生成 AI 利活用ガイドブック[38]（2024 年/経済産業省）

2024 年 7 月に発行された本文書は、生成 AI を利用したコンテンツ制作の企画・検討や、利用する生成 AI サービスの選択、リーガルチェック、さらに生成 AI の利用に関する社内ガイドラインの作成などへの活用が推奨される。

4) テキスト生成 AI 利活用におけるリスクへの対策ガイドブック（α 版）[39]（2024 年/デジタル庁）

2024 年 6 月に発行された本文書には、広島 AI プロセスや AI 事業者ガイドラインを踏まえ、行政業務において生成 AI を用いる際の指針として、生成 AI の中でもテキスト生成 AI とその利活用に焦点を当て、政府調達において考慮する点について、より具体的なリスクと対策が記載されている。

5) デジタル庁「DS-310 政府情報システムにおけるクラウドサービスの適切な利用に係る基本方針」[40]

″本方針の抜本的な改正を実施した 2022 年から２年が経過し、クラウドを取り巻く環境は更なる変遷を遂げている。クラウドはその黎明期においては「雲の向こうに隠蔽される仮想的なサーバー群」であったが、それが「インフラやソフトウェアを所有しない IT の形態」と進化し、今日では「スピーディかつ合理的に業務を進めるための手段」と、より包括的に捉えるべき存在に変遷している。例えば、オンデマンド・セルフサービスは単にすぐに使えるというだけではない。従来の個別価格交渉や納期という概念を不要にするものである。また、事前に利用量を予測して安全マージンを含めたリソースを当初から契約する必要もない。クラウドでは、全体の資源に余裕があれば、必要に応じて必要な量の資源を利用することができる。従来のオンプレミス環境では、個別価格交渉や納期、そして変化に対応しにくいシステム構成の

---

38 https://www.meti.go.jp/policy/mono_info_service/contents/aiguidebook.html

39 https://www.digital.go.jp/resources/generalitve-ai-guidebook

40 https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/e2a06143-ed29-4f1d-9c31-0f06fca67afc/a612d406/20250619_resources_standard_guidelines_guideline_08.pdf

ために、常に余裕を持ったリソースを確保しておく必要があった。しかしクラウドでは、そのような「安全マージン」の必要性は低くなる。同様に紙での報告、人手に頼った作業、過剰なテスト実施等、従前の習慣はクラウド環境への移行を契機に見直されることが望ましい。2025 年の改正においては、これらの変化も本方針に取り込んでいく。〞

6)行政の進化と革新のための生成 AI の調達・利活用に係るガイドライン（令和 7 年 5 月 27 日 デジタル社会推進会議幹事会決定）[41]

政府は様々な業務への生成 AI の利活用促進とリスク管理を表裏一体で進めるため、「行政の進化と革新のための生成 AI の調達・利活用に係るガイドライン」（令和 7 年 5 月 27 日 デジタル社会推進会議幹事会決定）を策定した。これは、政府における AI ガバナンスやベストプラクティスの共有体制、生成 AI の調達・利活用において留意すべきリスク等についての考え方、政府が利活用する生成 AI 全体の機能性や品質及び費用対効果の向上等について、AI 事業者ガイドラインや「政府機関等のサイバーセキュリティ対策のための統一基準群」等の既存のガイドライン及び諸外国政府のルールの動向等を踏まえ整理し、国の政府職員等向けのガイドラインとして示したものである。

7)医療デジタルデータの AI 研究開発等への利活用に係るガイドライン（2024 年/厚生労働省）[42]

民間企業等と共同で AI を活用した医療機器の研究開発等を実施するにあたり、個人情報保護法の下、医療機関等において診療で得られ、既に保管されている医療情報を利活用する際の法的・技術的な取扱いについて示している。医療情報は、一般的に、個人情報保護法上の個人データに該当する上、同法上の要配慮個人情報にも該当し、極めて機微な性質を有していることから、個人情報保護法で規定する仮名加工情報を作成し、運用するにあたって不可欠である実践的な指針として、また、医療情報の利活用に対する社会的な信頼の確保に貢献することを目的としている。

---

[41] https://www.digital.go.jp/resources/standard_guidelines

[42] https://www.mhlw.go.jp/content/001310044.pdf

8) 農業分野における AI・データに関する契約ガイドライン（2024 年/農林水産省）[43]

農業関係者等（農業関係者、農業指導普及員等）が持つノウハウ等を活用して、AI などを利用したシステムやサービスに反映させたり、あるいはノウハウを含むデータを用いて、システムやサービスを利用したりする場合の取決めに関する留意事項や条項例が示されている。また、スマート農業全般で用いられることが想定される農業関係者が持つデータについて、その利活用のために研究機関や企業等、プラットフォーム事業者などに提供する際の農業関係者の利益を配慮した条項例や留意事項が示されたものである。

9) 自治体における AI 活用・導入ガイドブック＜導入手順編＞（2025 年 12 月改訂）[44]

自治体向けの AI の利用方法や利用上の留意事項については、本ガイドラインを参考としつつ、自治体に特化した内容を記載した「自治体における AI 活用・導入ガイドブック＜導入手順編＞」（2025 年 12 月改訂）が公表されている。改訂後の本ガイドブックは、特に生成 AI の利活用により飛躍的な業務効率化が期待される点を、自治体における利活用事例とあわせて提示するとともに、ガバナンス確保のための体制構築、要機密情報の取扱い、人材育成の考え方等の留意事項についても提示している。

10) 防衛省「装備品等の研究開発における責任ある AI 適用ガイドライン」[45]

近年の人工知能（AI）技術の急速な発展と普及に伴い AI の性能が大幅に向上した一方で、新たな課題も浮上している。AI は、人間が定めた明確なルールや条件に基づいて動作するのではなく、与えられたデータからルールや知識を自ら学習し、それに基づいて新たな結果を出力する。このため、学習データの偏りなどに起因するバイアスや誤判断が生じる可能性など、AI 特有の技術的なリスクが存在する。また、AI の性能が向上したために AI が人間や社会に与える影響が大きくなっている。AI の推論過程や判断根拠等の不透明性、

---

[43] https://www.maff.go.jp/j/kanbo/tizai/brand/keiyaku.html
[44] https://www.soumu.go.jp/main_content/000820109.pdf
[45] https://www.mod.go.jp/atla/soubiseisaku/ai_guideline/ai_guideline_ver.01.pdf

個人情報保護やプライバシーへの懸念、著作権侵害の可能性、AI による偽情報の生成と拡散、AI 導入による雇用への影響など、技術的なリスクだけでなく AI 特有の倫理的・法的・社会的なリスクが存在する。これらの課題に対処するため、AI 技術の責任ある研究開発と利用に関する国際的な議論が活発化している。各国政府や国際機関、企業、学術団体などが、AI の倫理ガイドラインや規制フレームワークの策定に取り組んでいる。

防衛省は、防衛省・自衛隊の AI 活用に関する考え方を部内・部外に示すため、「防衛省 AI 活用推進基本方針（令和６年７月防衛省）」（以下「基本方針」という。）を策定した。基本方針の中では研究開発における防衛省・自衛隊独自のガイドラインを策定することとしている。

11)AI 利活用における民事責任の解釈適用に関する手引き（案）[46]

本書は AI 利活用における民事責任の在り方に関する研究会の議論を取りまとめるものであり、AI 利活用の場面についての不法行為法上の論点を中心に、現行法がどのように適用され得るかの方向性を示し、AI の開発・提供・利用に関わる当事者の予測可能性を高め、AI 利活用の推進及び損害発生時の円滑な解決に資することを目的とするものである。このような目的の下、各想定事例の解説においても、どのような要素が当事者に責任が生ずる可能性を高めるか、あるいは低めるかといった解釈適用の方向性を中心に記述するとともに、可能な限り広い事案の解決の参考となるような考え方を抽出し、また、それぞれの考え方が妥当するための前提や射程を示すよう試みている。留意を要する点として、個別具体的な事例に対し現行法がどのように適用されるかを最終的に判断する権限は裁判所に属しており、本書に記載された考え方が裁判所においてそのまま採用されることを保証し得るものではない。しかし、裁判例の蓄積を通じた解釈適用の明確化には一般に長期間を要するところ、有識者の議論に基づき、現時点で可能な限り合理的な考え方や判断が分かれるポイントを示すことで、本書が一つの法解釈の叩き台となり、以て新しいルール形成の一助となることを願っている。

---

46 https://www.meti.go.jp/shingikai/mono_info_service/ai_utilization_civil/004.html

12)AI のセキュリティ確保のための技術的対策に係るガイドライン[47]

総務省のサイバーセキュリティタスクフォースの下に設置された「AI セキュリティ分科会」において、AI システムの開発者及び提供者における、AI のセキュリティ確保に向けた技術的対策例を示すことを目的とした「AI のセキュリティ確保のための技術的対策に係るガイドライン」を 2026 年 3 月に策定。

１．６　民間の主な AI 関連ガイドライン

1)FUDA 生成 AI ガイドライン（2024 年/一般社団法人金融データ活用推進協会）[48]

生成 AI 独自の特徴とリスクの整理に加え、金融機関において生成 AI の利活用を推進するために実践的ルールを作る上で留意すべき AI 原則および法規制まで網羅的に取りまとめ解説されたものである。

2)ヘルスケア事業者のための生成 AI 活用ガイド[49]（2024 年/日本デジタルヘルス・アライアンス）

本ガイドにおいては、まずはヘルスケア領域で最も広く活用されていると考えられる文章（テキスト）生成 AI を対象とされる。なお、本ガイドは今後技術やサービスの進展を踏まえて随時アップデートを行う予定とされている。

3)医療・ヘルスケア分野における生成 AI 利用ガイドライン[50]（2024 年/非営利共益法人 医療 AI プラットフォーム技術研究組合）

生成 AI の医療現場における利用にともなうリスクと対策を示し、医療現場における生成 AI の導入と利用を促進することを目的としている。本ガイドラインは、医療機関・薬局等で生成 AI を利用する人、もしくは生成 AI の開発に携わる人を対象読者として想定されている。

---

47 https://www.soumu.go.jp/main_sosiki/kenkyu/cybersecurity_taskforce/index.html

48 https://www.fdua.org/activities/generativeai

49 https://www.jri.co.jp/page.jsp?id=107056

50 https://haip-cip.org/assets/documents/nr_20241002_02.pdf

4) ISO/IEC 42001

マネジメント標準としては、ISO/IEC42001 Information technology-Artificial intelligence-Management system が、監査組織の基準としては、ISO/IEC42006 Information technology - Artificial intelligence - Requirements for bodies providing audit and certification of artificial intelligence management systems 等を適宜参照することが有用である。

2 主要な組織、体制

1) AI 戦略本部[51]

2024 年 12 月 26 日に AI 制度研究会によって公表された『中間とりまとめ（案)』に対する議論を踏まえ、石破総理より全閣僚からなる『AI 戦略本部』の設置が指示された。各省庁や自治体において、インフラなどにおける AI の導入実態を把握し、ガイドラインの見直しなどの対応が進められていく。

2) 人工知能政策推進室 [52]

2025 年 8 月 1 日に、内閣府は人工知能関連技術の研究開発及び活用の推進に関する法律（AI 法）に規定される AI 戦略本部の事務局として、科学技術・イノベーション推進事務局に「人工知能政策推進室」を設置した。今後、AI 政策に係る行政各部の施策の統一を図るために必要となる事項の企画及び立案並びに総合調整に関する事務を担う。

3) AI セーフティ・インスティテュート[53]（AISI：エイシー）

2024 年 2 月に 10 府省、5 政府系機関が共同して AI セーフティ・インスティテュートを設立した。AISI は政府の AI セーフティに関する取り組みを支援する機関である。また、幅広く情報収集をしていることから、日本国内における AI セーフティに関する情報のハブとなる役割が期待されている。さらに AI セーフティに関しては国際的な取り組みが非常に速いスピードで進む中で、国際的な場への参加と貢献を期待されている。

---

[51] https://www.kantei.go.jp/jp/103/actions/202412/26ai.html
[52] https://www8.cao.go.jp/cstp/stmain/20250801ai.html
[53] https://aisi.go.jp/

4)GPAI 東京専門家支援センター[54]（GPAI：ジーペイ）

2024 年 7 月に GPAI(THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGNECE)東京専門家支援センターが設立された。AI に関する優先課題に取り組む産業界、市民社会、政府、国際機関、学界の専門家による研究やプロジェクトの運営・管理を支援する機関である。

GPAI 東京専門家支援センターは、SAFE プロジェクトの一環で、汎用 AI の安全性に関するツールキットを開発した。このツールキットは、世界の 500 を超える組織により検討されている汎用 AI に係るリスクと対策を分かりやすく詳細に整理しているのが特徴である。

[54] https://www2.nict.go.jp/gpai-tokyo-esc/