

AIガバナンスに資する透明性レポート ハンドブック（第1.0版）

巻頭言

AIをめぐる環境は、この数年で驚くほど変わりました。生成AIの普及により、私たちは日々、便利さと同時に「どう使うべきか」「どこまで任せてよいのか」といった問いに向き合うようになりました。いま、社会が求めているのは、技術を止めることなく、責任をもって活かし続けるための「AIガバナンス」です。

けれども、その実践は一人の担当者だけでは進められません。経営層、技術部門、法務、広報、人事、そして現場のスタッフまで、組織全体で考え、少しずつ形をつくっていくことが大切です。本ハンドブックは、そのための手がかりとして、G7/OECDの広島AIプロセス（HAIP）をもとに、透明性を高める報告の進め方を整理しました。

ここで紹介するステップやワークシートは、完璧な仕組みを作るためのマニュアルではありません。むしろ、自社の現状を見つめ、関係者と対話を始めるための出発点です。AIガバナンスを「書類」ではなく「文化」として根づかせる、その小さな一歩を後押しできれば幸いです。

2025年11月27日

江間有沙
工藤郁子
実積寿也

SUMMARY

AIの社会的活用が進む中で、リスク管理や説明責任、透明性の確保は国際的な共通課題となっています。

企業や組織は、技術をどう使うかだけでなく、どのように責任を持って使い続けるかを示すことが求められています。

本ハンドブックは、こうした背景を踏まえ、広島AIプロセス（Hiroshima AI Process: HAIP）の枠組みを通じて、AIガバナンスを実践的に進めるための手引きです。

前半の「概要編」では、AIガバナンスの必要性やHAIPの概要を紹介し、後半の「実践編」では、HAIPへの報告準備から提出までの具体的な手順を示します。

付属のワークシートは、自社の既存資料をもとに整理しながら記入できる設計で、初めての組織でも実践しやすくなっています。

本書全体を通じて、完璧な遵守よりも「現状を正直に示し、改善を重ねる」ことを重視し、AIガバナンスの実践と透明性の文化を組織に根づかせることを目的としています。

CONTENTS

ハンドブックとワークシートの目的と使い方

ハンドブック

概要編

経営層・広報・法務など、AIガバナンスや透明性報告の方針を検討・判断する立場の方に向けて、HAIP報告の意義、参加のメリットと全体像を解説します。

- ・第1章
AIガバナンスと透明性
- ・第2章
広島AIプロセス（HAIP）を理解する
- ・第3章
レポートの主な読者とその視点の可能性

実践編

各部署と連携しながら実際に情報を整理・作成を進める実務担当者に向けて、報告書作成の手順や社内調整の方法を具体的に示します。

- ・第4章
HAIP報告の全体像
- ・第5章
HAIP報告用ワークシート
- ・第6章
HAIPの今後の展望

ワークシート

実務担当者が各部署に回覧し、情報を記入・共有・集約するための作業用ツールとして活用できます。

設問	解答例	詳細	備考

概要編

なぜ透明性が求められているのか

AIの活用が社会のあらゆる領域に広がる中で、信頼・説明責任・透明性を確保することが国際的な課題となっています。

生成AIによる誤情報や偽情報の拡散、画像生成をめぐる著作権問題、アルゴリズムの偏りによる差別的な結果、そしてAIシステムの不具合やサイバー攻撃に起因する事故や損害。こうした「AI由来の重大インシデント」は、多様化し、組織はそれらにどう対応しているのかを示す責任が問われています。

多くの企業や機関はすでにAIガバナンスの体制を整備し、リスク管理や倫理指針を策定しています。次に求められているのは、その取り組みをどのように社会に伝え、理解と信頼を築くかという段階です。

透明性レポートや開示の仕組みは、単なる説明資料ではなく、組織が自らの価値観とリスク対応を外部と共有し、対話を通じて改善を重ねるためのツールです。

法令遵守だけでなく、企業の価値観や倫理を踏まえて「何を優先するか」「どこまで透明に開

示するか」を自ら決めることが、社会からの信頼の基盤になります。

AIの透明性レポートを公開することは、トラブルを防ぐための保険であると同時に、社内の意思決定を明確にし、リスクとイノベーションの両立を可能にする経営の道具でもあります。

AIをどう活用するかだけでなく、どう責任を持って使い続けるかを考える、それがいま、あらゆる組織に求められています。

Case 1 : NTT株式会社 技術企画部門 AIガバナンス室 根本 宗記

NTTでは、適切なAIの活用を推進し、AIリスクへの対応を統制するリーダーシップを確保するために、Co-CAIO (Co-Chief Artificial Intelligence Officer) を設置しています。

Co-CAIOの重要な役割であるAIリスクマネジメントは、競争領域ではなく協調領域と位置づけています。そのため、NTTグループが事業を展開する主要国や国際的な連携の中で、AIに関する法規制やガイドラインがどのように整備・運用されているかを的確に把握したうえで、NTTグループとしての方針を策定しました。

2023年の広島G7サミットでは、グローバルAIガバナンスが主要な議題の一つとして取り上げられ、「広島AIプロセス」が発表されました。これを契機に、NTTはOECDが主導する広島AIプロセスに基づいた報告枠組みを検討する国際タスクフォースに参加する機会を得ました。





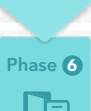


以降、広島AIプロセスやNTTグループのAIガバナンスに関して、日本および各国の政府関係者、議会、学術関係者との議論を行う機会が増えています。こうした取り組みはメディアにも取り上げられ、NTTグループはAIリスクマネジメントに先進的に取り組む企業として広く認識されるようになりました。さらに、主要な国際連携である広島AIプロセスへの積極的な貢献は、NTTグループ内のグローバル施策に対する求心力を高める効果をもたらしました。

今後も、広島AIプロセスの活用が一層進み、他者の権利を侵害しない“安心・安全で信頼できるAI”が世界に広がっていくことを期待しています。

AIガバナンス7つのフェーズ

本表は、組織におけるAIガバナンスを段階的に整備していくための7つのフェーズを示したものです。各段階の目的と関係部署を明確にすることで、自社の取り組み状況を整理し、今後の方向性を検討する際の参考としてください。

本ハンドブックでHAIP提出を考えている企業・組織の方々はフェーズ4ないし5までは整備されているという前提で進んでください。

フェーズ	何をするかの例	誰が関与するかの例
 Phase 1 情報収集と現状把握	AI の利用状況やリスクを把握し、既存のルール・対応を整理します。	技術部門、リスク管理部門
 Phase 2 ビジョンの策定	組織内の AI の活用目的と価値基準を明文化し、倫理指針や行動原則を整備します。	経営層、法務、広報
 Phase 3 組織体制の整備	責任の所在と意思決定プロセスを明確化し、有識者委員会等を設置します。	経営層、人事、コンプライアンス
 Phase 4 リスク評価と対策	偏り・誤情報・偽情報・セキュリティ・悪用などを AI 特有のリスクを分析し、必要な対策を講じます。	技術部門、リスク管理部門
 Phase 5 監査と評価	組織内の AI ガバナンス実践状況を定期的に評価し、課題を抽出・改善します。監査を通じて対応の妥当性や有効性を確認します。	経営層、監査部門
 Phase 6 透明性の確保と対話	AI の運用方針やリスク対応を社会一般やクライアント、利用者に説明します。情報公開だけでなく、説明方法を設計します。	広報、経営層、政策渉外担当
 Phase 7 人材育成と文化定着	倫理・技術・経営を横断する人材を育成し、AI ガバナンスを組織文化として根付かせます。	全社、教育担当部門

第2章 広島AIプロセス(HAIP)を理解する

広島AIプロセス（Hiroshima AI Process: HAIP）は、AIの開発と利用における透明性と説明責任を高めるための国際的な枠組みです。

本章では、HAIPが生まれた背景と意義、その基本構造、報告の仕組みと特徴と参加するメリットを紹介します。

なぜHAIPに注目するのか

AIの透明性や説明責任を求める動きは、欧州、北米、アジア諸国など世界中で広がっています。それぞれ独自の目的と手法を持ち、どれか一つが優れているわけではありません。

本ハンドブックでは、その中でも2023年のG7広島サミットを契機に始まった国際的な枠組みである広島AIプロセス（HAIP）を例に、透明性を高める実践的な方法を紹介します。

HAIPを取り上げる理由は三つあります。

① 国際的な公開性

G7合意に基づきOECDが運営するプラットフォームで各国の報告が公開され、比較・参照が可能なこと。

② 柔軟な仕組み

法的拘束力を持たず、国や業種を超えて自主的に参加できること。

③ 現実的な姿勢

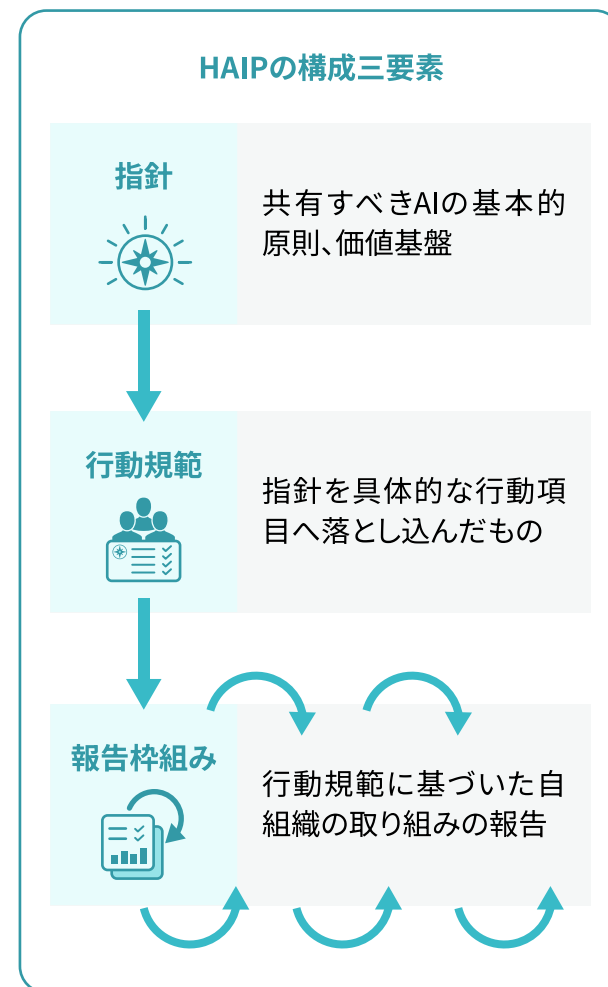
「完璧さ」ではなく、「現時点での取組状況を正直に開示すること」に価値を置くアプローチであること。

本書でHAIPを扱う目的は、特定の制度を推奨するのではなく、**透明性を通じてAIガバナンスを強化する実践モデル**として学ぶことにあります。

HAIPの構成要素

HAIPは、AIの信頼性向上と国際的な協調を目的とした**自主的なガバナンス枠組み**で、右記の三要素で構成されています。

これらは、国や制度の違いを超えて参加できるよう設計され、AIガバナンスの「協調と学習の場」として機能しています。



報告枠組みの仕組みと特徴

OECDのオンラインプラットフォームを通じて、HAIPに基づく報告は一般公開されています。

当初の対象は「高度なAIシステム開発者」でしたが、現在ではAI開発者だけではなくAI提供者も参加しています。

開発者はモデル設計やリスク評価に焦点を置き、提供者はサービス運用や利用者対応を中心に報告しています。

B2B企業は専門的記述を、B2C企業はわかりやすい説明を重視するなど、多様な表現形式が見られます。

HAIPは一律の形式を求めず、組織の状況や読者に応じた柔軟な開示を尊重しています。重要なのは「完全さ」ではなく、正直さとタイムリーな更新です。課題も含めて開示することで、相互学習と改善の循環が生まれ、透明性文化の形成に寄与しています。

参加組織一覧

2025年11月現在、24社の組織が報告枠組みに参加しています。

各国・地域別の参加企業をOECDへの報告提出順に並べました。

現在もいくつかの企業が新規提出に向けて準備をしています。

日本

1. KDDI Corporation
2. SoftBank Corp.
3. Preferred Networks
4. NEC Corporation
5. Fujitsu
6. Rakuten Group, Inc.
7. NTT（2025年9月に更新版提出）
8. Hitachi, Ltd.
9. ABEJA, Inc.

アメリカ

1. West Lake research & education service, a division of Palo Alto Research
2. Microsoft
3. Salesforce
4. Anthropic
5. OpenAI
6. Google
7. Amazon

その他の国

1. Data Privacy and AI（ドイツ）
2. KYP.ai GmbH（ドイツ）
3. TELUS（カナダ）
4. Fayston Preparatory School（韓国）
5. ai21（イスラエル）
6. MGOIT（ルーマニア）
7. TELUS Digital（カナダ）
8. Milestone（デンマーク）

HAIPに報告するメリット

以下は、実際にHAIPに提出をした企業に対して行ったヒアリングから得られた参加のメリットの一例です。

(1) 国際的な信頼・調達・投資への効果

HAIPへの参加は、調達や投資の信頼性を高める手段として注目されています。企業がAIガバナンスの取組状況を公開することで、国際的なパートナーや投資家からの信頼を得やすくなります。特に近年では、ESG投資の観点からもAIガバナンスへの関心が高まっており、透明性を確保することが投資判断に直結する重要な要素となっています。

(2) 中小企業・スタートアップへの実務的效果

HAIPは法的拘束力を持たない自主的な枠組みのため、リソースの限られた企業でも参加しやすい仕組みです。ISOのような監査手続きは不要で、必要最低限の情報を整理して報告できる点が実務的です。これにより、限られた体制でも国際的な信頼を示しやすくなり、新しい市場や投資機会へのアクセスを広げる効果があります。

(3) 社内ガバナンスとリスク管理の強化

HAIP報告は、外向けだけではなく実際の運用とのギャップに気づくなど、社内体制の整備やプロセスの改善にも役立ちます。方針や責任分担を明確にし、AI特有のリスクを可視化することで改善が進みます。年次更新を通じて、説明責任や倫理意識を重視する文化が組織に根づきます。

(4) 採用・顧客・社会的信頼への波及効果

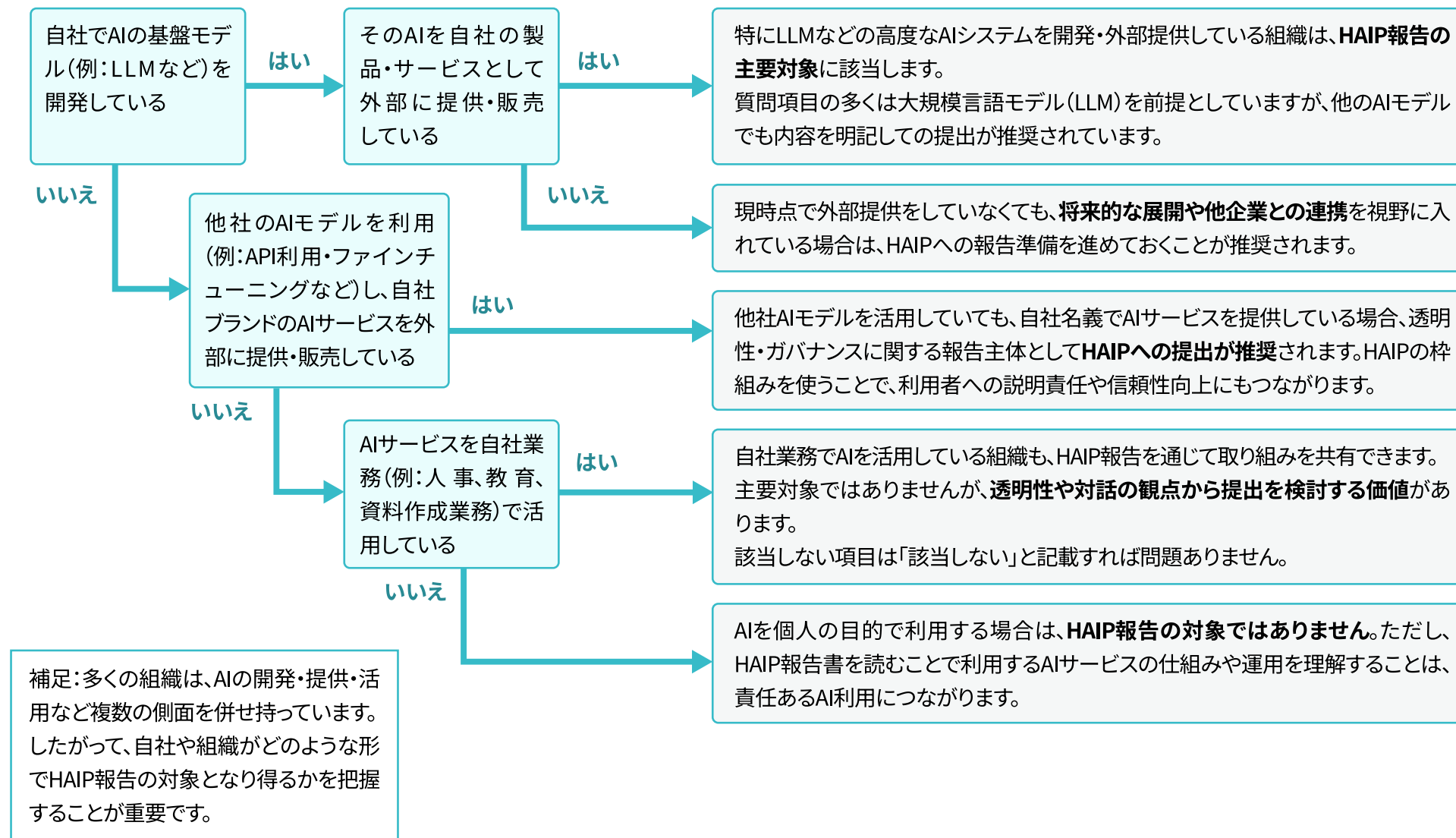
AIガバナンスへの取組は、採用や顧客との信頼構築にも直結します。近年、学生や若手エンジニアの間では、倫理的かつ責任ある企業姿勢への関心が高まっています。HAIP報告を通じて、自社の価値観や責任あるAI活用方針を明確に示すことは、優れた人材の獲得や、顧客・ビジネスパートナーに対する信頼向上につながります。

(5) 日本における整合性と実務的意義

日本では2025年にAI推進法が成立し、第十三条等でAIの透明性と説明責任を事業者の責務としています。政府は国際的な枠組みであるHAIPを指針運用に整合させる方針を示しており、企業がHAIPに参加することで、国内外双方での信頼と実効性を高めることができます。

HAIPにおけるAIとの関わり方と対象者分類

以下は「自分の組織はHAIPのどの対象区分にあたるか」を判断するためのフローチャートです。



全てのAI関係者向けの広島プロセス国際指針

「全てのAI関係者向けの国際指針」は、開発企業を対象にした11項目と、利用者に情報共有を求める最後の12項目目から成り立ちます。

1. AIライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度なAIシステムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる
2. 市場投入を含む導入後、脆弱性、及び必要に応じて悪用されたインシデントやパターンを特定し、緩和する
3. 高度なAIシステムの能力、限界、適切・不適切な使用領域を公表し、十分な透明性の確保を支援することで、アカウンタビリティの向上に貢献する
4. 産業界、政府、市民社会、学界を含む、高度なAIシステムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組む
5. 特に高度なAIシステム開発者に向けた、個人情報保護方針及び緩和策を含む、リスクベースのアプローチに基づくAIガバナンス及びリスク管理方針を策定し、実施し、開示する
6. AIのライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威に対する安全対策を含む、強固なセキュリティ管理に投資し、実施する
7. 技術的に可能な場合は、電子透かしやその他の技術等、ユーザーがAIが生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入する
8. 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先する
9. 世界の最大の課題、特に気候危機、世界保健、教育等（ただしこれらに限定されない）に対処するため、高度なAIシステムの開発を優先する
10. 国際的な技術規格の開発を推進し、適切な場合にはその採用を推進する
11. 適切なデータインプット対策を実施し、個人データ及び知的財産を保護する
12. 高度なAIシステムの信頼でき責任ある利用を促進し、貢献する

<https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document03.pdf>

高度なAIシステムを開発する組織向けの国際行動規範

行動規範は以下の11項目から成り立ちます。

1. AIライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度なAIシステムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる
2. 市場投入を含む導入後、脆弱性、及び必要に応じて悪用されたインシデントやパターンを特定し、緩和する
3. 高度なAIシステムの能力、限界、適切・不適切な使用領域を公表し、十分な透明性の確保を支援することで、アカウンタビリティの向上に貢献する
4. 産業界、政府、市民社会、学界を含む、高度なAIシステムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組む
5. 個人情報保護方針及び緩和策を含む、リスクベースのアプローチに基づくAIガバナンス及びリスク管理方針を策定し、実施し、開示する
6. AIのライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威に対する安全対策を含む、強固なセキュリティ管理に投資し、実施する
7. 技術的に可能な場合は、電子透かしやその他の技術等、ユーザーがAIが生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入する
8. 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先する
9. 世界の最大の課題、特に気候危機、世界保健、教育等（ただしこれらに限定されない）に対処するため、高度なAIシステムの開発を優先する
10. 国際的な技術規格の開発を推進し、適切な場合にはその採用を推進する
11. 適切なデータインプット対策を実施し、個人データ及び知的財産を保護する

<https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document05.pdf>

報告枠組み

HAIPの報告枠組みは7セクション、39項目から構成されます。

① リスクの特定と評価

(Risk Identification & Evaluation)

- AIシステムに関するリスクや脆弱性の特定方法
- インシデントや新興リスクへの対応手順
- レッドチーミングや外部評価の実施状況

② リスク管理と情報セキュリティ

(Risk Management & Information Security)

- データ品質とバイアス低減の取組み
- プライバシー保護、知的財産権の取扱い
- サイバー・物理的セキュリティ対策
- 脆弱性管理や内部不正対応の仕組み

③ 透明性

(Transparency Reporting)

- システムの能力・限界・適正／不適正用途の公開
- 学習データ、評価手法、モデル更新情報の開示
- プライバシーポリシーや利用者向け情報提供

④ 組織統治・インシデント対応・透明性

(Organizational Governance)

- AIガバナンスを組織統治に統合する方法
- スタッフ教育・トレーニング体制
- インシデント発生時の記録と情報共有手順

⑤ コンテンツ認証と来歴

(Content Authentication & Provenance)

- AI生成物の識別方法(ラベリング、ウォーターマーク等)
- 国際標準に基づく来歴情報の付与

⑥ 研究・投資

(Research & Investment)

- 安全性・信頼性・公平性に関する研究投資
- コンテンツ真正性や来歴管理に向けた協働取組み
- 社会的リスク緩和への研究投資

⑦ 公益の促進

(Advancing Human & Global Interests)

- AIを通じた社会的・環境的便益の最大化
- ユーザー教育やデジタルリテラシー支援
- 市民社会との協働による社会課題解決

<https://transparency.oecd.ai/instructions>

第3章 レポートの主な読者とその視点の可能性

レポートの想定読者は多様です。それぞれの読者がどのような立場からHAIP報告書を読む可能性があるのかについて整理しました。

政策担当者

政策担当者は、AIガバナンスやリスク管理の現状を把握するために透明性レポートを活用します。そこから得られる知見は、新しい制度やガイドラインの検討に役立ちます。

また、企業が重点を置く分野や課題を把握することで、支援や政策の重点領域を判断する材料にもなります。レポートは優劣を競うものではなく、政策との補完関係を見いだす資料として読むことが大切です。

利用者

利用者にとって透明性レポートは、自分が利用するサービスがどのような考え方や仕組みで運用されているかを知る手がかりです。

「何ができるのか」「どんなリスクがあるのか」を理解することで、より適切な利用判断ができます。また、企業の説明責任を確認し、意見や対話を始めるための窓口にもなります。

投資家

投資家は、透明性レポートから企業のAIガバナンス体制を把握します。完璧さよりも、課題を含め正直に開示しているかという姿勢が重視されます。

透明性そのものが信頼の指標となり、ESG評価や投資判断の重要な要素になります。継続的に更新される報告を追うことで、企業の改善力やリスク対応力も見極められます。

顧客企業

顧客企業にとって透明性レポートはベンダーのデューデリジェンスやサードパーティリスク管理の出発点となります。AIシステムのリスク対策、運用体制が開示されることで、取引先の信頼性を客観的に評価できます。

また、継続的に更新される報告を参照することで、パートナー企業の改善状況やリスク対応力を把握し、長期的な協働の判断材料として活用できます。

その他関係者

透明性レポートは、以下のような多様な立場でも活用されています。

同業他社：業界として協調して取り組むべき内容を、他社の取組を参照して理解し、自社の改善やベンチマークに活用。

研究者・専門家：AIガバナンスの実態分析や政策提言に利用。

NGO・市民団体：社会的影響への取組を確認し、対話や提言に活用。

国際機関：各国の報告を比較し、制度の調和や協力を促進。

メディア：企業の取組や課題を市民に伝える際の一次情報源として活用。

透明性レポートは、社会の多様な立場の人たちにとって対話と改善の出発点です。報告を作成する組織にとっても、多様な読者を意識することが、誠実で伝わる情報発信につながります。

読み手が重視するポイントと報告作成のヒント

レポートの読者が重視するポイントも多様です。それぞれの読者が重視するだろうポイントと報告作成のヒントについて整理しました。

順位づけではなく信頼を重視

透明性レポートは、企業を点数化したり順位づけしたりするためのものではありません。

読み手は「どのような背景で」「どんな姿勢で」AIガバナンスを進めているのかを見ています。

報告の中で「まだ実施していない」「課題が残っている」と記すことも、誠実な情報開示として評価されます。

文脈を踏まえて理解する

報告を読む際、読み手は項目ごとに企業の重点を見ます。

たとえば、「どのリスクを重視しているか」「セキュリティ対策の成熟度」「利用者への説明の範囲」などです。

また、年ごとの変化を見ることで、改善や方針転換の方向性を読み取ります。

このような観点を意識して報告を作成することで、読み手が理解しやすく、組織の信頼を得やすい内容になります。

開示の姿勢そのものがメッセージになる

報告の細部よりも、「何をどこまで説明する姿勢を持っているか」が信頼性の判断基準になります。

AIガバナンスの成熟度は、形式的な完璧さよりも、「対話を続ける意思」「社会に説明する意欲」によって測られるといえます。

この意味で、透明性レポートは社会との信頼を築くためのコミュニケーションツールです。

Case 2 : KDDI株式会社 先端技術企画本部 高木幸一

KDDIがHAIP 報告枠組みは参加したねらいは、ビジョンに掲げる「誰もが思いを実現できる社会」をつくるために信頼あるAIの提供が不可欠であり、同取組への参画がその実現に重要になると考えたためです。具体的には、対外的にはAIガバナンス活動の透明性を高めステークホルダーや社会に対し信頼を訴求すること、社内的には自社の取り組みや進捗状況を整理し客観的に振り返るための機会とすることであり、これらの実施に貢献したものと考えています。

報告の提出に当たっては本ハンドブックの実践編に記載のとおり社内関係者への照会と責任者の承認を経っていますが、最大の工夫点は、回答をドラフトするにあたり生成AIを活用したことです。公開情報に即し回答するというHAIP報告枠組みの方針を踏まえ、生成AIに質問票の質問を投じることで、公開情報をベースとした客観的な回答案を数秒で作成しました。もちろん、生成AIの出力を踏まえ、人手による最終確認と判断は必須でしたが、機密情報の懸念を回避しつつ、ベースのドラフトを作成するためには非常に有益でした。

詳しくは下記リンクに記載していますので覗いてみてください。

<https://tech-note.kddi.com/n/ne47aa01787a0>

実践編

第4章 広島AIプロセス(HAIP)報告の全体像

ここからは、実際に報告書を作成・提出する担当者の方に向けた「具体的な手順とポイント」をまとめます。
HAIPの報告は難しい専門作業ではなく、準備と工夫をすればどの組織でも取り組めるように設計されています。
大事なのは、完璧な遵守ではなく、透明性をもって現状を示すことです。

取りまとめ担当の役割とスキル

HAIP報告の推進には、部門横断的に情報を整理し、経営層と現場をつなぐ「取りまとめ担当/部署」の役割が重要です。

求められるのは、以下のようなスキルです。

1. 倫理・法務・技術を統合して課題を整理する**分析力**
2. 関連部門に協力を仰ぎ、意見をまとめて合意形成を促す**調整力**
3. 専門性とわかりやすさを両立して記述する**コミュニケーション力**
4. 英訳後の意味を確認できる**言語リテラシー**
5. 正直な開示と改善を重ねる**誠実さと継続性**

取りまとめ担当は、AIガバナンスを社内に根付かせる**対話と信頼形成の推進役**です。

報告までのステップ

OECDウェブサイトの質問項目、あるいは本ハンドブックのワークシートなどを見ながら、何が求められているかを確認します。OECDのウェブサイトでは、すでに提出している企業の報告を読むことができます。

本ハンドブックでは、報告までのステップの一例を示します。次のページから具体的に見ていきましょう。

OECDのウェブサイトから報告をしたり、報告書を一覧で見ることができます。2025年11月時点で、8か国24組織が報告を提出し、一般に閲覧可能な形で開示されています。



HAIP Reporting Framework (OECD.AI) : <https://transparency.oecd.ai/>



STEP1 担当・チーム体制の整備

推進する担当やチーム体制には会社の規模や社内ガバナンスの進み具合によりいろいろな方法が考えられます。

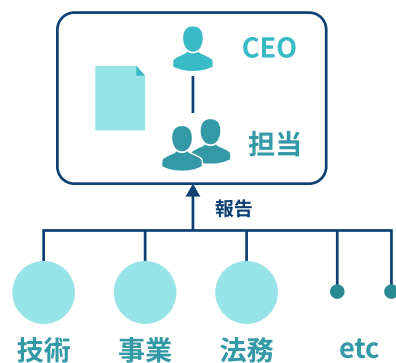
初めは小規模な取り組みでも、継続的なHAIP報告作業を通じて、ガバナンス体制をより強化できるよう徐々に体制を整備していくことが重要といえます。

専任担当者型 (小規模体制)

CEOおよびその直下の1～2人のメンバーで執筆する。

兼任で技術、事業、法務など網羅的に報告内容をカバーする。

少人数のスピード重視型。

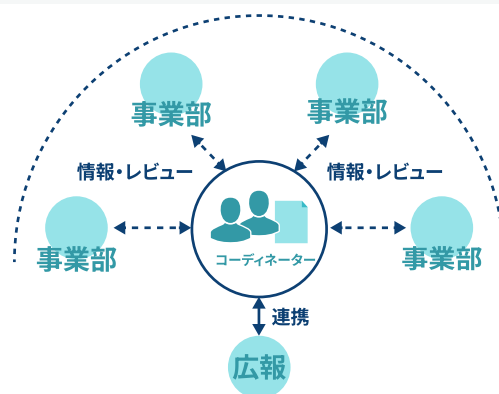


調整役主導型 (プロジェクトベース)

一人または少人数のコーディネーターが各部署に働きかけ、情報を収集する。
例えば渉外担当やガバナンス担当が窓口、コーディネーター役となる。

ドラフトは少人数で作成、各部署がレビューで参加、広報とも連携する。

情報が幅広く網羅でき全社的認識が高まる一方、調整工数が多い。

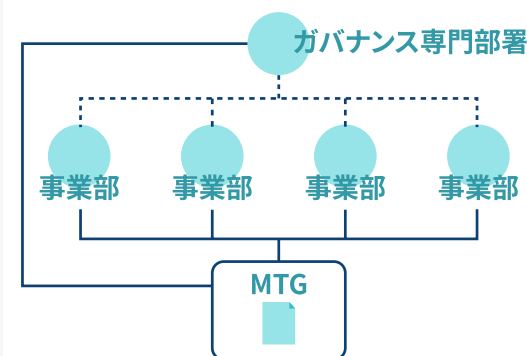


部門横断型 (ガバナンス専門部署主導)

既存のガバナンス専門部署等がコーディネーターとなって、関連部署が共同で作業する。

各部署が参加する会議体等で進める。

体制によってはリソース負担が大きいが、継続的な取り組みに展開できる。



STEP2 既存資料の特定と収集、スコープ決定

資料収集

既存の社内資料（AI利用ガイドライン、リスク管理規程、透明性レポート、CSR/ESG・統合報告書など）で参照できるものを収集します。

外部で公開している資料（Webサイト、プレスリリース、技術ブログなど）も確認します。公開資料は添付することでゼロから作る必要はありません。足りない情報はどの部署に問い合わせるかを特定します。

スコープの決定

HAIPは、AI開発者、AI提供者、大企業から中小企業、研究機関に至るまで、幅広い組織が参加できる柔軟な枠組みです。

この柔軟性はHAIPの大きな特徴であり、同時に、自社は

① どの立場として

② 誰に向けて

③ 何を報告するのか

を自ら明確に定めることが重要になります。

① どの立場として

HAIPはもともと**高度なAIシステムの開発者**を主な対象として設計されました。しかし、実際にはサービス提供者や公共機関など、多様な立場の組織が報告を行っています。

企業によっては開発者でもありかつ提供者であるという場合もあり、どの立場で報告を書くかとも考慮に入れる必要があります。

AIシステム開発者を対象にした設問項目でもあるため、現状では提供者（アプリの開発者等）や研究機関では答えにくい設問があることも事実です。その場合は、「本システムでは該当しない」や「アプリ開発の場合は、以下の対策をとっている」などと設問をどのように解釈したかを書く必要もでてくるでしょう。

今後多くの企業が提供していくことによって設問項目自体の改善も行われることが期待されています。

② 誰に向けて

報告書の想定読者も多様です。すでに提出した企業へのインタビューからは、国際機関、政府関係者、企業パートナー、市民など多様な想定読者に言及されました。

提出企業が想定している企業や組織以外の人も読む人もいる可能性もあるため、できるだけ平易な書き方をすることが望ましいですが、例えばB2B企業は既存の資料がすでにB2B向けということもあるでしょう。

ビジネスパートナー向けに専門用語が使われることが多い報告書になる可能性もあります。

そのため報告書の想定読者は誰であるかを冒頭で明示することが、読者の理解を助けるうえで有効となります。

③ 何を報告するか

報告の対象（どの単位で書くか）も明確にしておく必要があります。
たとえば以下のように設定できます。

モデル単位

特定のAIモデル（例：大規模言語モデル、画像生成モデル）やシステムに関する説明

サービス単位

製品・サービスの運用と利用者対応に関する説明

組織単位

全社的な方針、ガバナンス体制、リスク管理の枠組みに関する説明

これらのいずれを対象とするかを最初に明示しておく、読者は報告の範囲と意図を把握しやすくなり、透明性の効果が高まります。

例えば、該当項目に応じて、「サービス単位＋組織全体」の二層で記載していくことで、対象がより明確になりわかりやすい開示となります。

HAIPは形式が定められすぎていないからこそ、こうした「スコープの自律的な定義」が重要な第一歩となります。

Case 3：マイクロソフト社 Amanda Craig Deckard、Kate Purchase、Hector de Rivoire、小島治樹

マイクロソフトは、広島AIプロセス報告枠組みに、責任あるAIに関する国際的な共通規範の形成を推進し、運用段階における透明性の在り方を示すことを目的として参加しました。本取組は、当社が掲げるグローバル・ガバナンスの強化および産業界における連携強化と軌を一にするものです。

準備に際しては、Frontier Governance Framework、Responsible AI Transparency report、Responsible AI Standardなどの社内リソースを活用し、法務・技術・政策各部門が連携して一貫性と明確性を確保しました。また、技術的な深度と分かりやすさの両立のために複数回にわたる厳密な部門横断レビューを実施しました。

対外的には、透明性や説明責任の具体的な仕組みについて他の事業者や政策当局と意見交換を行う貴重な機会となり、AI原則をガバナンスや開発プロセスに実装するための実務的知見を共有することができました。

今後は、モデル開発者のみならず、アプリケーション開発者や提供者を対象に範囲を拡大することで、AIライフサイクル全体にわたる透明性を一層向上できるものと考えます。各主体に応じたモジュール型の報告構造を導入することで、より幅広い参加を促進し、報告内容の比較可能性を高め、エコシステム全体における説明責任と信頼性の強化につながることを期待しています。

STEP3 関係部署と会議体の計画、承認プロセスの特定

各項目について、足りない情報、協議が必要な情報と関係部署、承認プロセスを特定し、必要な会議をスケジュールします。下表では HAIP 報告書の各項目について、主にどの部署が中心となって記述・整理を進められるかの目安を示しています。例えば、ドラフト作業までは関係部署内で作成、承認し、レビューを部署横断の会議で行うことも考えられます。以下、HAIP 報告枠組みの7セクションごとに関連する社内業務の例を示します。

HAIP 報告項目	関係部署	主な役割・関連業務
① リスクの特定と評価	経営企画／法務／情報システム／品質管理／ R&D ／内部監査／サービス設計・プロダクト部門、モデル開発部門	全社リスク統制、法的リスク評価、脆弱性特定、出力品質確認、技術的リスク把握、監査チェック
② リスク管理と情報セキュリティ	情報セキュリティ／品質管理／データガバナンス／法務／サービス設計・プロダクト部門、モデル開発部門	運用・サイバーセキュリティ、データ品質管理、知財・個人情報保護
③ 透明性	経営企画／ R&D ／広報・IR・渉外／サービス企画／サービス設計・プロダクト部門、モデル開発部門	方針統制、学習データ開示、利用者・投資家・政策担当者等対外的なコミュニケーション
④ 組織統治・インシデント対応	経営企画／人事・研修／リスク管理／内部監査／サービス設計・プロダクト部門、モデル開発部門	教育設計、インシデント対応体制、報告手順確認、監査
⑤ コンテンツ認証と来歴	R&D ／情報システム／法務／サービス設計・プロダクト部門、モデル開発部門	ウォーターマーク開発、来歴管理システム実装、規制確認
⑥ 研究・投資	経営企画／ R&D ／サステナビリティ／外部連携／サービス設計・プロダクト部門、モデル開発部門	安全性研究、社会的リスク低減、共同研究推進
⑦ 公益の促進	経営企画／サステナビリティ／サービス企画／人事／サービス設計・プロダクト部門、モデル開発部門	社会的価値創出、市民社会連携、従業員教育

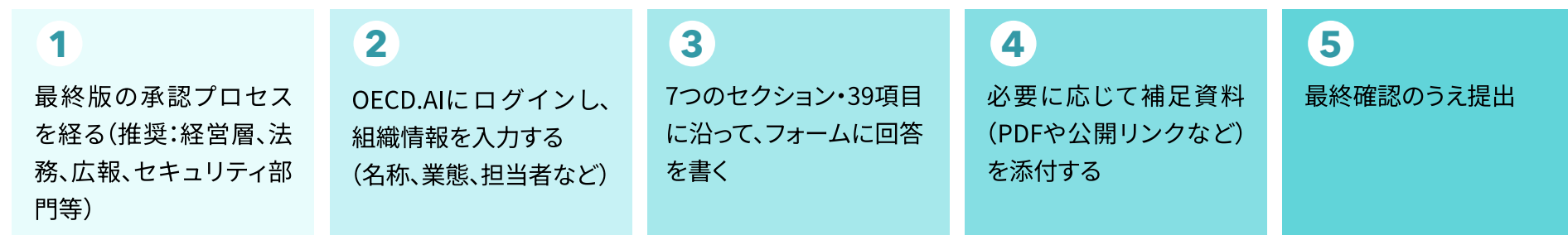
STEP4 社内の合意形成（会議等の開催）と承認

報告に向けた合意形成のプロセスも各組織の体制によりますが、関係部署から担当を招集して報告の目的を共有しながら進めることで、社内の意思疎通や社内体制の強化につながります。会議開催による進行ステップの例を示します。回数、アジェンダと執筆期間は参考目安としてください。

会議	アジェンダ
第1回（キックオフ）	<ol style="list-style-type: none">1. HAIP報告の背景と目的の共有2. 報告対象範囲（サービス／モデル／全社）とスコープの確認3. 項目ごとの執筆担当者の確認4. 未担当項目を関連部署に執筆依頼・調整
2-3週間程度で、各担当者が公開文書をベースに資料確認・執筆	
第2回（情報整理）	<ol style="list-style-type: none">1. 項目ごとに執筆内容の確認2. 執筆にあたって課題が出てきた場合、議論 議論例1) 機密性情報以外で新たに公開できる事項があるかの精査 議論例2) 機密性情報であるため「取り組みをしていないように見られる」ことを回避するための対応策の議論
2-3週間程度で、担当者が資料確認・執筆	
第3回（ドラフトレビュー）	<ol style="list-style-type: none">1. 各セクションの初稿を共有、確認2. 情報公開の範囲や読みやすさの表現を全体確認
1-2週間程度で修正、最終ドラフトを日本語と英語で執筆	
第4回（最終承認準備）	<ol style="list-style-type: none">1. 日本語と英語の最終ドラフトの確認2. 提出手続き、社内外への公表方針の確認

STEP5 OECD提出

報告はすべてOECD.AIの透明性プラットフォームからオンラインで行います。流れは次の通りです。



STEP6 継続的改善と年次更新

OECD.AIサイトで公開されて閲覧可能になったあと、年に1回の更新が原則です。前年の提出月を基準にしたサイクルが推奨されています。

1年以上更新がない場合、担当者に連絡があり、それでも更新がなければウェブサイトからは削除される予定です。

更新時には「去年から何を改善したか」を明示することで、外部にも努力が伝わりやすくなります。

Case 4 :株式会社Preferred Networks エンジニアリングマネージャー 大野健太

弊社は大規模言語モデル（LLM）の開発を手がける一方、AI面接など、AIガバナンスが特に重要となる領域でもプロダクトを提供しています。HAIP報告枠組みへの参加は、対外的な信頼獲得と社内体制整備の両面で大きな効果がありました。

第一に、お客様からの信頼獲得です。AIリスク管理やガバナンス体制を重視されるお客様に対

し、国際的な取り組みであるHAIP報告枠組みに基づき、我々の取り組みを具体的に報告できることは大きな強みとなります。

第二に、LLM開発者としての意見反映です。LLM開発企業は資本力のある大企業が多い中、我々のような比較的小規模なデベロッパーの意見や実情を、ガバナンスに関する議論の場へ反映させる貴重な機会となりました。

社内での大きく二つのメリットもありました。一つは、ビジネスメンバーがお客様へガバナンスの取り組みを説明するための道具が整理されたこ

とです。これにより、お客様とコミュニケーションがスムーズとできるようになりました。

もう一つは、HAIPへの取り組みに関する公表情報を社内でも共有することで、「会社としてガバナンスに取り組まなければならない」という意識を全社的に醸成する一助となったと考えています。

HAIPへの参加は、我々のガバナンス体制を客観的に示し、ステークホルダーとの対話を促進する有効な手段となりました。

第5章 HAIP報告用ワークシート

本ハンドブックでは、HAIP報告の全体像を把握しながら、報告作業がなるべくスムーズに行えるよう、報告作業用ワークシートを作成しました。

ワークシートの見方

- ワークシートはすべての業態および規模の組織を対象としています。
- それぞれの質問項目ごとに、多くの企業の共通内容をなりうる記載例と、さらに詳細な記載例を示しています。参考にしながら、自組織の取り組み状況に応じて公開内容を検討するようにしてください。
- なお、ここでの記載例は、すでにHAIPに報告している企業の回答事例（2025年4月提出の報告レポート）を参考に要約・抽象化しているものです。そのため、各文例は一つの正解を示すものではありません。また、今後も参加企業や関係各所の意見を取り入れながら、アップデートしていく予定です。
- 自組織の回答を作成する際には、これらの記載例を参考にしつつ、自組織の現状に照らして回答するようにしてください。

ワークシートは以下リンクからダウンロードできます
https://www.tc.u-tokyo.ac.jp/ai1ec_event/16044/



ワークシートのイメージ

質問項目	回答例（要約）	回答例（詳細）	コメント	評価項目	回答例	回答例	回答例	回答例
組織の概要（名称・業態・規模）	「〇〇株式会社」の概要を要約して記載している。	「〇〇株式会社」の概要を要約して記載している。また、業態・規模についても記載している。	「〇〇株式会社」の概要を要約して記載している。また、業態・規模についても記載している。	組織の概要	〇〇株式会社	業態・規模	業態・規模	業態・規模
組織の取り組み（取り組みの概要・取り組みの成果）	「〇〇株式会社」の取り組みの概要を要約して記載している。	「〇〇株式会社」の取り組みの概要を要約して記載している。また、取り組みの成果についても記載している。	「〇〇株式会社」の取り組みの概要を要約して記載している。また、取り組みの成果についても記載している。	組織の取り組み	〇〇株式会社	取り組みの概要	取り組みの成果	取り組みの成果
組織の取り組み（取り組みの概要・取り組みの成果）	「〇〇株式会社」の取り組みの概要を要約して記載している。	「〇〇株式会社」の取り組みの概要を要約して記載している。また、取り組みの成果についても記載している。	「〇〇株式会社」の取り組みの概要を要約して記載している。また、取り組みの成果についても記載している。	組織の取り組み	〇〇株式会社	取り組みの概要	取り組みの成果	取り組みの成果
組織の取り組み（取り組みの概要・取り組みの成果）	「〇〇株式会社」の取り組みの概要を要約して記載している。	「〇〇株式会社」の取り組みの概要を要約して記載している。また、取り組みの成果についても記載している。	「〇〇株式会社」の取り組みの概要を要約して記載している。また、取り組みの成果についても記載している。	組織の取り組み	〇〇株式会社	取り組みの概要	取り組みの成果	取り組みの成果

回答の書き方：基本とポイント

レポートの読者が重視するポイントも多様です。それぞれの読者が重視するだろうポイントと報告作成のヒントについて整理しました。

1. 読みやすく、伝わる記述を意識する

回答は、想定する読者（例：一般利用者、取引先、規制当局など）を明確にし、そのリテラシーに合わせたわかりやすい表現を心がけましょう。

例：「利用者」について主に誰が利用者になるかを明記しつつ、その対象に向けた対策について主に記載（セクション4.Cやセクション5等）する。

B2BやB2Cなど事業形態に応じ、専門用語は括弧書きで補足し、概要を本文に記載します。リンクのみを貼るのではなく、主な内容やポイントを文章で説明することが望ましいです。

例：主要な公開文書へのリンクを示したうえで、その中で特に強調すべき項目を本文で要約する。

2. 未実施項目も正直に記載する

未実施の項目を空欄にせず、「現時点では未実施」「今後検討予定」と明記し、背景や課題、検討スケジュールを簡潔に示しましょう。

例：「〇〇については未実施。来年度中に対応方針を策定予定。」

Case 5：ソフトバンク株式会社 AIガバナンス推進室長 浦野憲二

ソフトバンク株式会社は「情報革命で人々を幸せに」という理念のもと、AIの社会実装と倫理的活用の両立を目指しています。その一環として、行政・企業・教育機関が協働し、国際的にAIの信頼性や倫理を議論・実践する「広島AIプロセス（HAIP）」の趣旨に賛同し、2024年度より広島AIプロセス・フレンズグループ・パートナーズコミュニティに参画しました。生成AIをはじめ急速に進化する技術環境の中で、透明性・公平性・説明責任を確保しながら社会的信頼を築くことが不可欠との認識によるものです。

参加に際しては、AIガバナンス推進室を中心に関連部門が連携し、HAIP提出文書の作成・検証を実施。行動指針と

3. 実践における留意点

報告は「完璧さ」ではなく「現状と課題の開示」が基本です。「実施済み」「検討中」「未実施」を区分して整理し、今後のロードマップを添えると透明性が高まります。

営業秘密や競争上の機微情報は事前に整理し、開示可能な範囲で説明します。

また、報告書冒頭には報告のスコープ（モデル単位／サービス単位／組織全体）を明記しましょう。

中小企業やスタートアップなど、限られたリソースの組織では、ワークシートの基本項目を確実にカバーし、必要に応じて補足情報を加える形でも十分です。

自社AI倫理原則の整合性を確認しつつ、企業として果たすべき責任を明確化しました。議論を通じてAI倫理への理解が深化し、社内意識の共有も進みました。

HAIP参画により、「責任あるAI企業」としての信頼向上や官民連携の強化、さらにAI倫理教育や生成AI運用ルールの改善など、社内外で実質的な成果を得ています。

第6章 HAIPの今後の展望とOECDへの提案

HAIP（広島AIプロセス）は、AIガバナンスの国際的な透明性確保の枠組みとして大きな意義を持ちます。しかし、制度が実際に運用される中で、いくつかの実務的な課題や改善の余地も明らかになってきました。本章では、現場での利用経験や技術動向を踏まえ、今後の改訂に向けた方向性を提案します。

(1) 報告内容の構造整理と重複の解消

現在のHAIP報告書の設問には、内容が重なったり、同様の情報を別の項目で再度求められたりする箇所があります。これらを統合・整理することで、より効率的で理解しやすい報告フォーマットになると考えられます。

特に、実務担当者の負担を軽減し、回答の一貫性を保つためには、重複項目の統合や設問の再編成が有効です。

(2) 開発者と提供者に応じた柔軟な設計

現行のHAIPでは、LLM（大規模言語モデル）を開発する企業と、そのモデルを利用してサービスを提供する企業が、同一の設問フォーマットで回答することが求められています。しかし、両者の役割や責任範囲は大きく異なります。

今後は、開発者用テンプレートと提供者用テンプレートを分ける、あるいは共通部分と選択的回答部分を明確化するなど、より実態に即した柔軟な設計が望まれます。

(3) 報告データの閲覧・比較・更新の仕組み

現状、OECDのウェブプラットフォーム上では、報告がPDF形式で掲載されており、検索・比較・経年変化の追跡が難しいという課題があります。また更新をすると以前のバージョンを見ることができません。そこで具体的には以下のような改善が考えられます。

- 項目単位での**一覧表示・検索機能**の追加
- **バージョン履歴**を残し、過去の報告との変更点を可視化
- **差分報告項目**の導入（更新時に「何を変更したか」を記載）

これにより、利用者が企業の改善の軌跡を把握しやすくなり、社会全体での「学習と改善のエコシステム」が形成されます。

Case 6：楽天グループ株式会社 情報セキュリティ・プライバシーガバナンス部 財津健次

楽天グループは、急速に変化するAI環境に適応していく社会に貢献するため、「G7広島AIプロセス」に基づく運用枠組みに参加し、透明性レポートを作成しました。レポート作成の過程で、AIガバナンスを構成する主要分野の関係者が相互に活動内容を理解し、社内外で利用可能な共通言語や共通見解を認識することができました。今後も本レポートの継続的な更新を通じて、誠実な情報開示に努めてまいります。

本ハンドブックは、江間有沙（東京大学）、工藤郁子（大阪大学）、実積寿也（中央大学）を中心としてタスクフォースを作り、議論を重ねて作成したものです。とりまとめと窓口を東京大学江間研究室が行いました。これに対して国内外の研究者、OECDをはじめとする国際機関、すでにHAIP報告を行った企業、総務省をはじめとする関係省庁、AIセーフティ・インスティテュート（AISI）、GPAI東京専門家支援センター、日本ディープラーニングセンター（JDLA）などの機関から協力を得て作成しました。

また本ハンドブックは、東京大学の運営費交付金による支援および一般社団法人日本ディープラーニング協会（JDLA）の助成を受けて作成しました。

本ハンドブックを効果的に利用するために推奨されるリテラシー水準は、**日本ディープラーニング協会（JDLA）のG検定合格レベル**を想定しています。

G検定相当のリテラシーとは

- AIの基本概念（機械学習、深層学習、生成AI）を理解していること
 - AIの社会的影響や倫理的課題について基礎知識を持つこと
 - ビジネスや社会課題にAIを適用する際のリスクや留意点を把握できること
- この水準があれば、本ワークシートの質問意図や用語の意味を理解でき、実務的な記入が可能になります。

学習リソース例

- JDLA公式教材（AI For Everyone、日本語解説書籍）
- OECD.AIポータルの学習コンテンツ

HAIPの内容に関するフィードバック

時間や所属組織の関係上、お名前の記載が叶いませんでしたが、有益なフィードバックを寄せていただいた方々がいることも付け加え、フィードバックを寄せていただいた方々に感謝申し上げます。

足立光司（NTTドコモソリューションズ株式会社）
荒堀淳一（富士通株式会社）
伊藤彰嗣（楽天グループ株式会社）
伊藤淳史（株式会社日立製作所）
大野健太（株式会社Preferred Networks*）
小田淳司（KDDI株式会社）
工藤真一（ソフトバンク株式会社）
小島治樹（日本マイクロソフト株式会社）
財津健次（楽天グループ株式会社）
佐保優一（ソフトバンク株式会社）
高木幸一（KDDI株式会社）
高橋正和（株式会社Preferred Networks）
永沼美保（日本電気株式会社）
根本宗記（NTT株式会社）
原山優子（GPAI東京専門家支援センター）
人見彰彦（株式会社セールスフォース・ジャパン）
古川直裕（株式会社ABEJA）
細野桃子（富士通株式会社）
村上明子（AIセーフティ・インスティテュート）
吉永和弘（日本電気株式会社）
Alik Foinikopoulou（Salesforce, Inc.）
Merve Hickok（Center for AI and Digital Policy）

*2025年HAIP報告時点の所属
（五十音順、敬称略）

企画・制作補助

池田直子・岡田恵美（東京大学）
大谷光（日本ディープラーニング協会/東京大学）
狩野愛歌（フリーランス）
英訳者：川本麻衣子
デザイン制作：渡邊 一也

AI技術の進化や利用形態の変化に伴い、何を報告するのかや評価基準も継続的に見直す必要があります。

とりわけ、生成AIやマルチモーダルAIの登場、倫理・安全性・社会影響に関する国際的な議論の深化など、環境は急速に変化しています。HAIPはこれらの動向を踏まえ、年次的な更新とステークホルダー参加型の改訂プロセスを確立することが期待されます。

その際には、企業・研究機関・政策担当者・市民社会など、実際に利用や報告に関わる多様な立場の意見を反映させることが不可欠です。

HAIPは、透明性を通じて信頼を築くための国際的な試みとして大きな成果を上げつつあります。本ハンドブックが示すように、その価値は「完璧な遵守」ではなく、「進化し続ける仕組み」にあります。

今後は、報告内容・技術環境・公開方法の三位一体での改訂を進めることで、より開かれた、持続可能なAIガバナンスの実現が期待されます。

本ハンドブックも、皆様からのフィードバックや、報告枠組み自体の改定を踏まえて、バージョンを更新していく予定です。

また、2025年11月27日にはハンドブックを踏まえた透明性レポートの在り方に関する公開イベントも開催予定です。イベントはウェブサイトで動画や報告書も公開していく予定ですので、ご覧ください。その他関連イベントも以下のウェブサイトにとまっています。

<https://sites.google.com/g.ecc.u-tokyo.ac.jp/ema/projects/ai-governance>

本ハンドブックは、クリエイティブ・コモンズ・ライセンス「CC BY 4.0」の下で提供されています。このライセンスにより、出典を明記すれば、誰でも自由に共有・改変・翻訳が可能です。

引用・転載の際は、次の情報を明記してください。

江間有沙・工藤郁子・実積寿也（2025）

『AIガバナンスに資する透明性レポートハンドブック（第1.0版）』

東京大学東京カレッジ 江間研究室

<https://sites.google.com/g.ecc.u-tokyo.ac.jp/ema/projects/ai-governance>

ライセンス : CC BY 4.0



発行元・問い合わせ

東京大学国際高等研究所東京カレッジ

江間研究室

Mail: tg-event@tc.u-tokyo.ac.jp（研究室スタッフメールアドレス）

第1.0版 2025年11月27日