

Interpol/UNICRI

# 法執行機関における 責任あるAIイノベーションのための ツールキット

Toolkit for Responsible AI Innovation  
in Law Enforcement

## 日本語解説資料

November 2023



2023 年 11 月 20 日

## 法執行機関の AI システム導入に向けた取組みと AI ガバナンス： Interpol/UNICRI ツールキットの紹介と日本への示唆

江間有沙・東京大学国際高等研究所東京カレッジ  
宇田川尚子・東北大学法学研究科公共政策大学院  
尾崎愛美・筑波大学ビジネスサイエンス系  
後藤大・晴海パートナーズ法律事務所弁護士  
渡邊道生穂・AI 法研究会事務局長弁護士  
濱口海慧・東京大学公共政策大学院  
白石圭佑・東京大学公共政策大学院

### 1. はじめに：法執行機関における AI リスク評価

人工知能（AI）の活用が様々な場面で進展する中、警察を始めとする法執行機関が AI を活用する上では、AI システムのリスクを評価した上で、捜査の必要性と人権を侵害するリスクとのバランスをとることが重要と考えられる。このような観点から国際刑事警察機構（以下、インターポール）と国連地域間犯罪司法研究所（以下、UNICRI）は、AI 導入に当たって注視すべき責任ある AI 原則を提唱し、これを実装するための「法執行機関における責任ある AI イノベーションのためのツールキット（Toolkit for Responsible AI Innovation in Law Enforcement）」を共同で 2023 年 6 月に公表した<sup>1</sup>。

本ツールキットは、法執行機関が AI 導入にあたって考慮すべき項目を例示している。ただし、これらはあくまで参考情報という位置付けであり、各国・地域の組織構造や慣習、AI 利用等の実情に合わせてカスタマイズして検討を行っていくことが推奨されている。

AI ガバナンスに対する国際的な関心の高まりを踏まえれば、今後日本の法執行機関にもリスク評価の議論が及ぶことが予想される。実際、AI の開発・利用にあたっては、AI 開発者や利用者自らがリスクを評価しガバナンス機能を発揮することが重要であるとされている<sup>2</sup>。

---

<sup>1</sup> 本ツールキットはインターポール（<https://www.interpol.int/How-we-work/Innovation/Artificial-Intelligence-Toolkit>）と UNICRI（<https://unicri.it/topics/Toolkit-Responsible-AI-for-Law-Enforcement-INTERPOL-UNICRI>）の両サイトで公開されている。

<sup>2</sup> 日本の統合イノベーション戦略推進会議の下で開催されている AI 戦略会議が 2023 年にまとめた論点整理（[https://www8.cao.go.jp/cstp/ai/ai\\_senryaku/2kai/ronten.pdf](https://www8.cao.go.jp/cstp/ai/ai_senryaku/2kai/ronten.pdf)）において記述されている。実際リスク評価やガバナンスの方法論については、日本国内においては、経済産業省が公開している「AI 原則実践のためのガバナンス・ガイドライン ver. 1.1」

（<https://www.meti.go.jp/press/2021/01/20220125001/20220124003.html>）があり、2023 年現在、経済産業省と総務省によって新 AI 事業者ガイドラインの検討が進められている。また法執行機関ではない

このような状況を踏まえ、本稿は考え方の一つの指針として続く2章でAIツールキット公開の背景と全体構造を概説する。また、3章では日本の警察におけるAIシステム利用の状況事例をいくつか紹介したのち、本ツールキットを展開していくうえでの可能性と課題についての執筆メンバーによる見解を述べる<sup>3</sup>。

## 2. AIツールキット

### 2.1 インターポールとUNICRI

法執行機関におけるAI利活用に関する議論は、インターポールとUNICRIが共同で2018年ごろに立ち上げた研究グループに端を発する。この研究グループの成果は、「法執行機関のためのAIとロボティクス」（2019年）<sup>4</sup>、「責任あるイノベーション：法執行機関のためのAI」（2020年）<sup>5</sup>の報告書として結実しているほか、UNICRI単独で2020年にAIの利活用に関する事例集<sup>6</sup>を公開している。

また、2020年11月にはインターポールとUNICRIの主催で5日間にわたるオンライン

---

が、自治体が導入するAIリスク評価の実践例としては、加古川市が一般社団法人日本データマネジメント・コンソーシアム（JDMC）と行った「高度見守りカメラ」の事前倫理検証がある（<https://www.city.kakogawa.lg.jp/soshikikarasagasu/shiminbu/shiminseikatsuanshinka/ICT/38711.html>）。そのほか、東京大学AIガバナンスプロジェクトが展開しているリスクチェーンモデルのケース例として、裁判所における再犯可能性の検証AIが公開されている（[https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\\_Case06\\_Verification-of-Recidivism-Possibility-AI\\_JP.pdf](https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel_Case06_Verification-of-Recidivism-Possibility-AI_JP.pdf)）。

<sup>3</sup> 本稿は執筆メンバーの一人である江間（東京大学）の呼びかけにより、2023年8月より定期的な研究会にて議論をもとに執筆されたものである。なお、本稿で紹介している内容の一部は2023年11月20日に東京大学東京カレッジで開催されるウェビナー「法執行機関のAIシステム導入に向けた取組みとAIガバナンス：Interpol/UNICRIツールキットの紹介と日本への示唆」の参考資料としてウェブサイト（[https://www.tc.u-tokyo.ac.jp/ai1ec\\_event/10775/](https://www.tc.u-tokyo.ac.jp/ai1ec_event/10775/)）に公開している。また同ウェブサイトでは、インターポールとUNICRIの本ツールキット関係者によるビデオメッセージによる紹介も閲覧することができる（<https://youtu.be/PHFJHwtvFJc>）。

<sup>4</sup> Artificial Intelligence and Robotics for Law Enforcement, Interpol and UNICRI, 2019,

[http://www.unicri.it/in\\_focus/on/interpol\\_unicri\\_report\\_ai](http://www.unicri.it/in_focus/on/interpol_unicri_report_ai)

<sup>5</sup> Towards Responsible AI Innovation, Second Interpol-UNICRI Report on Artificial Intelligence for Law Enforcement, Interpol and UNICRI, 2020, <http://www.unicri.it/towards-responsible-artificial-intelligence-innovation>

<sup>6</sup> Special Collection on Artificial Intelligence, UNICRI, 2020, <http://www.unicri.it/node/3228>。事例集ではすでにAIを利用している国による以下の8事例が紹介されている。(1)刑法分野でAIを規制するツールとしてのバランシングテスト（アルゼンチン）、(2)警察における顔認証技術利用の公平性、信頼性（オーストリア）、(3)警察のAIの目的外利用に対するシステム不安への対策としての目的制限設計（ベルギー）、(4)証拠から証明へ：イタリア司法裁判所における社会ネットワーク分析（イタリア）、(5)ヘルスケアにおけるAI：リスク評価と刑法（イタリア）、(6)犯罪司法システムへのAIからの挑戦例としてのディープ・フェイク（デンマーク）、(7)性目的のための人身売買へ対抗するAI駆動型分析（オランダ）、データ政権：政府がデータを規制する方法を理解する分析的ガイド（アメリカ）

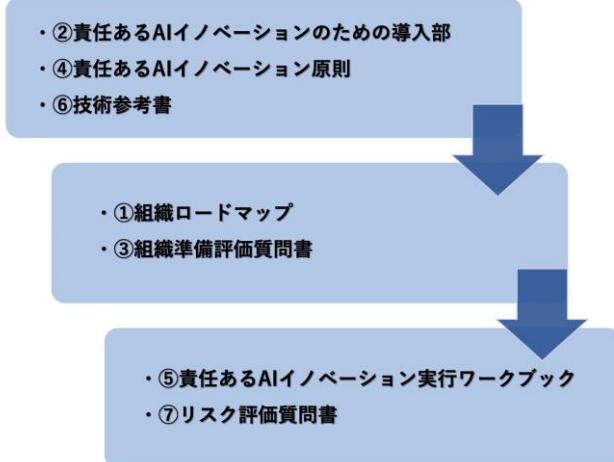
会議が開催され<sup>7</sup>、80か国から600人以上が集まり、AIの悪用・誤用、オンラインでの児童虐待やテロリストによるSNS使用への対抗としての法執行機関におけるAI利用、民間における法執行のためのAI開発、刑事司法制度におけるAIの使用、AIと刑事責任、AIとローンの相互作用などが議論された。会議では、法執行機関がAIを利活用していくことが不可欠であり、そのためにもAIの透明性、アカウンタビリティや信頼性が重要となることが確認された。また、今後も具体的なケーススタディを共有していくことや、ツールキットを共同開発していくことが提案された。

その後、欧州連合からの資金援助を受け、2020年11月よりインターポールイノベーションセンターとUNICRIの人工知能とロボット工学センターの指揮のもとツールキット開発が開始された。ツールを議論するにあたっては法執行機関、産業界、学術界、司法機関、警察官、人権の専門家などによる国際的な専門家グループが結成され、定期的なオンライン会合を経て、2023年6月に現行のバージョンが公開された。今後は2023年秋まで本ツールキットに関するフィードバックを募集し、ツールキットバージョン2としてウェブベースでの公開や、政府機関や各国法執行機関へのトレーニングプログラムへと展開が目指されている<sup>8</sup>。

## 2.2 ツールキットの全体構造

本AIツールキットは、法執行活動に対するAIシステムの導入に関心のある法執行機関が、責任あるAIイノベーションを理解し、実際にそれぞれの業務に取り入れることを支援することを目的としている。「法執行機関」とは、警察機能行使する警察及び類似の国家機関を念頭に置いている。ただし、本ツールキットは、法執行機関に限らず、刑事司法の当事者である裁判官、検察及び弁護士並びに民間セクター又は学術界の技術開発者及び市民社会など、幅広い利害関係者において理解されるべきであるとされている。

全体構造としては、本ツールキットは次の8つの文書から構成されている。すなわち、READMEファイルに続き、①組織ロードマップ、②責任あるAIイノベーションのための導入部、③組織準備評価質問書、④責



<sup>7</sup> Artificial Intelligence and law enforcement: challenges and opportunities. 2020, <https://www.interpol.int/News-and-Events/News/2020/Artificial-Intelligence-and-law-enforcement-challenges-and-opportunities>

<sup>8</sup> 今までの経緯や今後の展開については、以下を参照：<https://unicri.it/sites/default/files/2023-02/Factsheet%20AI%20Innovation%20LE.pdf>

任ある AI イノベーション原則、⑤責任ある AI イノベーション実行ワークブック、⑥技術参考書、⑦リスク評価質問書の 8 つである。

README ファイルは、①、②及び④を「Primary Guidance Document」、③、⑤及び⑦を「Practical Tool」、⑥を「Supporting Document」として分類する。そしてツールキット使用者に対し、まず②、④及び⑥を使用して基礎的な理解を習得し、次に①及び③を使用して各組織の自己評価を行なった後、実際に AI システムを開発、運用する際に⑤及び⑦を使用することを推奨している（図 1）。

以下本稿においては、内容が関連するものをまとめて取り上げることとし、責任ある AI イノベーションに向けた組織体制の整備とその評価に関する①及び③、その具体的な実行に関する④及び⑤、そして個別の AI システムのリスク評価に関する⑦に分けて紹介する。

なお、本稿において詳細な紹介は行わないが、②及び⑥においては、法執行機関における責任ある AI イノベーションに関する技術的、法的な概念及び用語の解説や、考慮されなければならない価値及び問題の説明などが行われている。法執行機関が実際に AI イノベーションを行う準備段階はもちろん、AI イノベーションが実際に行われている最中においても、適宜参照されることが望ましいと思われる。

### 2.2.1 責任ある AI イノベーションのための組織体制

本ツールキットは、責任ある AI イノベーションを実現するためには、法執行機関が組織面における準備を整えている必要があるとの考えに立脚している。このため、AI システムを導入する組織に必要であると考えられる構成要素等が組織ロードマップで示されるとともに、法執行機関の準備状況を評価するため組織準備評価質問書が用意されている。

その内容を概観するに、各組織には「組織文化」、「人材と専門性」及び「プロセス」の 3 要素が必要であるとされているところ、その中でも「組織文化」が主要な部分を占めるとされている。これは、AI システムの使用が人権と自由を尊重する形で行われているか等は、法執行機関に対する国民の認識や受止め、ひいては社会的信頼に影響を与えるものであるからであると説明されている。

この「組織文化」を実現するには、「正しいマインドセット」として 8 つの要素が必要であるとされている。すなわち、まず基本的な心構えとして①「一歩ずつ進めること」、②「『なぜ』から始めること」及び③「リスクに精通すること」の 3 点が挙げられている。次に、④「責任あるイノベーションにインセンティブを与えること」、⑤「新たな制度的アーキテクチャの必要性を認識すること」及び⑥「リソースの投入に備えること」という、組織の管理・運営に関する要諦が 3 点述べられている。そして、最後に⑦「反発に備えること」及び⑧「後退（撤回）に備えること」として再び 2 点の心構えが挙げられており、AI システムの停止や導入撤回までを選択肢に含めたマインドセットの重要性が強調されている。

次に、「人材と専門性」として、AI システムを活用する法執行機関には、AI 技術のみならず、倫理的かつ人権適合的な AI 使用に関する知見を有する人材と専門性が必要であると

されている。具体的には、①AI やデータ管理に関する「技術的コンピテンシー」、②刑事司法制度等に関する「領域コンピテンシー」、③法及び倫理に関する「ガバナンスコンピテンシー」及び④広報に関する「社会文化コンピテンシー」の4つの要素を挙げ、ポジションごとに必要なスキルの程度を示している。ただし、本ツールキットはこれら全ての専門家を組織内に抱えることまでは要求しておらず、必要に応じて外部機関との連携や商品・サービスを利用することでもよいとされている。

最後の「プロセス」は、法執行機関が、設計に始まるAIシステムのライフサイクル全体を通じて責任あるAIイノベーションを行えるようにするため、警察の長及び上級幹部等が推進すべき具体的な取組み等を指すものである。具体的なグッドプラクティスとして、①「組織全体のニーズ・能力評価の実施」、②「パブリックエンゲージメントのための基盤作り」、③「ガバナンスアプローチの定義」、④「リスクマネジメントポリシーの設定と遵守」及び⑤「責任あるAIイノベーション監督委員会の設立」が定められている。その内容を概観すると、まずはAIを導入すべきか、そうであれば何が必要かを評価し（①）、利害関係者との事前協議を開始し（②）、当該組織のAI利用のガバナンスのための戦略とリスク管理方針を策定し（③及び④）、ガバナンスアプローチの柱となる監督委員会を設立する（⑤）というものである。

そして、これらの準備状況については、**組織準備評価質問書**による評価が用意されている。同質問書は、上記3つの観点に関する各20問程度の質問から構成されており、「よく当てはまる」から「全く当てはまらない」までの5つの選択肢から1つを選ぶ形式である。例えば、「人材と専門性」に関する第1問は、「組織においてAIシステムの設計、開発及び使用に関与する幹部および職員は、『責任あるAIイノベーション原則』の十分な理解及び知識を有している。」（図2）といったものであり、責任あるAIイノベーション原則それ自体に対する理解と遵守を重視していることがうかがわれる。

| People and Expertise   |   |   |   |   |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
|--|---|---|---|---|--|---|--|--|--|--|--|---|---|---|---|---|--|---|---|---|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <p>This section of the assessment focuses on the individuals that play a key role in building and managing an agency's AI system, as well as those involved in interpreting the system's results. It is intended to help assess your agency's readiness in terms of its technical and non-technical expertise and competencies.</p> <p>▶ LEARN MORE ABOUT THE NECESSARY PEOPLE AND EXPERTISE IN THE ORGANIZATIONAL ROADMAP.</p>  |   |   |   |   |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| <table border="1"> <thead> <tr> <th colspan="6">Assign a score between 0 and 4 to reflect the extent to which you believe the statements below to be true for your agency</th> </tr> <tr> <th>0</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> </tr> <tr> <td>Not true; Have not considered this; Don't know</td> <td>Slightly true; We are starting to consider this</td> <td>Largely true; We are starting to implement this</td> <td>True; We have started to implement this</td> <td>Very true; We consistently implement this</td> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> </tbody> </table> |   |   |   |   |  | Assign a score between 0 and 4 to reflect the extent to which you believe the statements below to be true for your agency |  |  |  |  |  | 0 | 1 | 2 | 3 | 4 | Not true; Have not considered this; Don't know | Slightly true; We are starting to consider this | Largely true; We are starting to implement this | True; We have started to implement this | Very true; We consistently implement this | <input type="checkbox"/> |
| Assign a score between 0 and 4 to reflect the extent to which you believe the statements below to be true for your agency  |   |   |   |   |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| 0  | 1   | 2   | 3                                       | 4   |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| Not true; Have not considered this; Don't know   | Slightly true; We are starting to consider this | Largely true; We are starting to implement this | True; We have started to implement this | Very true; We consistently implement this |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| <input type="checkbox"/>   | <input type="checkbox"/>                        | <input type="checkbox"/>                        | <input type="checkbox"/>                | <input type="checkbox"/>                  |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| <input type="checkbox"/>   | <input type="checkbox"/>                        | <input type="checkbox"/>                        | <input type="checkbox"/>                | <input type="checkbox"/>                  |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| <p>1. The management and personnel in our agency involved in the design, development and use of AI systems have good understanding and knowledge of the <a href="#">Principles for Responsible AI Innovation</a>.</p>  |   |   |   |   |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |
| <p>2. The management and personnel involved in the design, development and use of AI systems follow the <a href="#">Principles for Responsible AI Innovation</a> and identify and mitigate any legal and ethical concerns and negative consequences of the use of AI systems in line with the Principles.</p>  |   |   |   |   |  |   |  |  |  |  |  |   |   |   |   |   |  |   |   |   |   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                          |

（図2）：「人材と専門性」第1問、第2問

これらの質問への回答に応じて0から4までのスコアが与えられ、その合計値により観点ごとにレベル5を最高、レベル1を最低とする5段階の「組織準備レベル」に振り分けられる。おおむねレベル5の組織では責任あるAIの理想的な活用が行われており、レベル3であれば適切に準備された本番稼働の初期段階程度、レベル1ではほとんど知見を有していないという程度の意味合いである。

## 2.2.2 責任ある AI イノベーションの実行

次に、法執行活動及び刑事司法制度に関する既存の確立された原則に基づく AI に関する考え方の枠組みに加え、法執行の観点から責任ある AI イノベーションをどのように実現できるかを法執行機関に示すことを目的として作成されたのが、**責任ある AI イノベーション原則**である。

同原則は 5 つのコア原則、すなわち①合法性、②危害の最小化、③人間の自律性、④公平性及び⑤グッドガバナンスから成り、それぞれにコア原則を補完する手段的原則が掲げられている(図3)。この 5 原則を総合すると、法執行における責任ある AI イノベーションを構成するのは、危害を最小限にとどめ、人間の自律性を尊重し、優れたガバナンスを支えとする合法的かつ公平な方法で、AI システムを開発、調達、使用することであるということができる。

これを踏まえ、それぞれのコア原則について概観する。

**合法性**は、法執行機関と AI システムとの関わりが合法的である必要があるというものであり、手段的原則として正当性、必要性、比例性の 3 つが掲げられている。

**危害の最小化**は、AI の開発、調達、使用に伴って生じ得る個人やコミュニティに対する危害のリスクを防止、排除、及び軽減することであり、手段的原則として堅牢性と安全性、正確性、人間と環境の健全性及び効率性が掲げられている。

**人間の自律性**とは、人間に影響を与えるいかなる決定も、最終的には人間によって行われることを求めるものであり、手段的原則として人的管理と監督、人間の主体性、プライバシー及び透明性と説明可能性が掲げられている。

**公平性**は、法執行機関が、AI システムとの関与の中で個人およびグループを公正かつ無差別に扱うこと及びより平等な社会の実現に貢献すべきことを意味しており、手段的原則として平等と無差別、社会的弱者の保護、多様性とアクセシビリティ及び係争性と救済が掲げられている。

**グッドガバナンス**は、法執行機関が、監査とアカウンタビリティのための組織的な構造を確立し、また、責任ある AI イノベーションの文化の推進を目指すべきであることを意味しており、手段的原則としてトレーサビリティと監査能力及びアカウンタビリティが掲げられている。

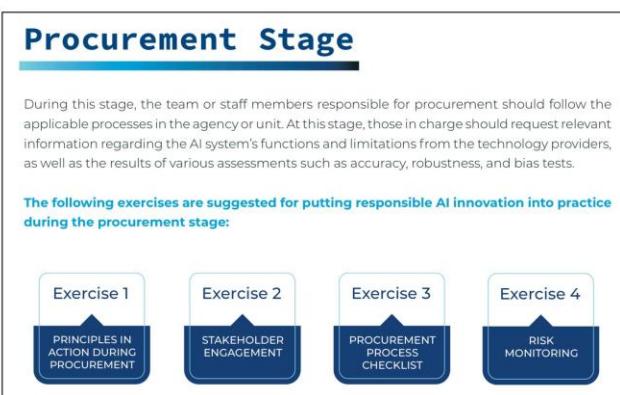
そして、以上に述べた原則を実践するためには、法執行機関が原則を理解し適用すること、利害関係者を特定し関与させること、結果を確認し必要に応じてこのプロセスを繰り返すことが有用であるとされる。かかる実践の具体的な枠組みは、**責任ある AI イノベーション実行ワークブック**において具体化されている。

そこで概観するに、同ワークブックは、法執行機関が責任ある AI イノベーション原則に沿った AI イノベーションを実行・維持するために必要な枠組みを提示するものであり、①計画段階、②設計・開発段階、③調達段階及び④使用・監視段階という AI ライフサイクルの各段階で確認すべき事項を示すとともに、「演習」として具体的な実践を示している。

**計画段階**では、AI システムの必要性と範囲、AI システムを入手するために利用可能な選択肢などについて議論すべきであるとされ、演習として、計画段階中の行動原則、利害関係者の関与、ギャップとニーズの分析、バリューマッピング、AI システムの入手方法の決定、ユースケース分析、誤用又は不正使用の可能性の特定及びリスク対応の 8 つが掲げられている。

**設計・開発段階**では、責任ある AI イノベーション原則を満たすように、最適なアルゴリズムアーキテクチャを選択し、トレーニングデータを特定及び収集し、それらのデータを適切に代表するモデルを設定すべきであるとされ、演習として、設計・開発段階中の行動原則、利害関係者の関与、データチェック、設計・開発における重要なポイント及びリスクの監視の 5 つが掲げられている。

**調達段階**は、法執行機関が AI システムの設計・開発を行わずに外部から調達することとした場合に適用するものであり、外部から提供される AI システムの機能、制限に関連する情報に加え、正確性、堅牢性、バイアステストなどの様々な評価結果について確認する必要があるとされる。そして、演習として、調達段階中の行動原則、利害関係者の関与、調達プロセスのチェックリスト、リスクの監視という 4 つが掲げられている(図 4)。



(図 4)：調達段階における演習

**使用・監視段階**では、AI システムが意図した通り使用されていることだけでなく、エラーやバグ等の包括的な使用状況を監視し、必要に応じて修正するべきであるとされ、演習として、使用・監視段階において法執行機関は、使用・監視段階中の行動原則、利害関係者の関与、展開プロセスのチェックリスト及びリスクの監視の 4 つが掲げられている。

### 2.2.3 AI システムのリスクアセスメント

**リスク評価質問書**は、法執行機関全体のリスク管理プロセスの一環として、**責任ある AI イノベーション原則**を遵守しない場合に個人及びコミュニティに生じる悪影響について、個別の AI システムに係るリスクを特定し、及び評価することを目的としたものである。

本ツールキットの特徴として、各リスクについてそれぞれ 5 段階の「可能性」と「影響度」のスコアを掛け合わせた「リスクスコア」が算出され、リスクの定量的な把握が可能

になっていることが挙げられる。これは、法執行機関は、AI システムにより生じる悪影響に対して限られたリソースでの対応を迫られることから、特定されたリスクに優先順位を付けて対応できるようにするためである。

具体的な質問は、「可能性」と「影響度」において分けて記載されており、24問の質問から構成されている。例えば、第1問は、「そのAI システムに関連する活動が、適用される法律又は規制に対する違反に繋がる可能性はどの程度あるか。」（可能性）、「こうした違反による悪影響は、個人及びコミュニティに対してどの程度深刻か。」（影響度）といったものであり（図5）、影響度に係る質問文は、全ての質問において大きく異なるものではない。

なお、同一又は類似の領域に係るリスクについて複数の質問が設定されているものもみられ、バイアスによる差別等に関するリスクについては特に多くの質問数が割かれている。最後に、本評価は定期的に行うべきものとされており、その結果はリスクを軽減するための措置などに活用されることとなるが、かかる措置が講じられた後にも再評価を行うことが求められている。

### 3. 日本における AI システムの評価の展開と課題

AI ガバナンスに対する国際的な関心の高まりを踏まえれば、今後日本の法執行機関にもリスク評価の議論が及ぶことが予想される。実際、AI の開発・利用にあたっては、AI 開発者や利用者自らがリスクを評価しガバナンス機能を発揮することが重要であるとされている。

一方で、本ツールキットは2章で紹介してきたように、特に組織体制の準備状況や法執行機関に限った AI ガバナンスの原則などに焦点をあてているのが特徴でもある。本稿では、このように法執行機関、とりわけ警察機構に焦点を当てている本ツールキットの可能性と同時に日本の法執行機関における AI の利活用に向けた方向性を検討する。

#### 3.1 警察等における AI 利用の実例

日本の公務部門においては、厳しい財政状況を背景として厳格な定員管理を迫られており、業務能率化を目的としたいわゆる行政 DX が求められている状況にある。そこで、政府全体の方針として「政府機関における AI の導入促進に向けた推進体制の強化と、それに

| Risk Assessment Questionnaire   |   |                  |            |              |  |
|---|---|------------------|------------|--------------|--|
| LIKELIHOOD (A)  | IMPACT (B)  | RISK LEVEL (AxB) | RISK OWNER | OBSERVATIONS |  |
| 1 - Very Unlikely; 2 - Unlikely; 3 - Possible; 4 - Likely; 5 - Very Likely  | 1 - Insignificant; 2 - Limited; 3 - Moderate; 4 - Severe; 5 - Catastrophic                        |                  |            |              |  |
| 1. What is the likelihood that activities related to the AI system will lead to non-compliance with applicable laws and regulations?* | How serious would the negative impact of such non-compliance be on individuals and communities? * |                  |            |              |  |

(図5)：リスク評価質問書 第1問

よる行政機能の強化・改善」<sup>9</sup>を図ることが謳われている。

このことは警察においても同様であり、警察庁は「人口減少等により限られた人的・財政的資源の下で最大限の効果を上げることができるよう、AI（中略）といった先端技術等を積極的に活用することを通じて警察力強化を図ることが求められている」<sup>10</sup>との認識を示している<sup>11</sup>。

日本の警察による AI システムの活用については、2019 年に警察庁情報通信局情報通信企画課に先端技術導入企画室<sup>12</sup>が設置されており、同室のリーダーシップの下、本格的な導入に先立つ実証実験が行われている段階にあるものが多い。先駆的な例として、マネーロンダリングやテロ資金の供与等に関するおそれがあるとして金融機関から届出がなされた「疑わしい取引」の年間 40 万件以上に上るデータを解析し、職員が注意を払うべき優先順位を付すというシステムがある。令和元年度から行われた同システムの実証実験において、システムの出力が従前行われていた職員の手作業による整理・分析の結果と齟齬がないことが確認されたため、令和 4 年から本格運用が開始されている。このほか、防犯カメラ映像から車種を判別するシステムや、SNS 上の書き込みにテキスト分析を行い規制薬物に関する情報を抽出するシステムに係る実証実験が行われている<sup>13</sup>。加えて、SNS に関しては、2021 年に警察庁が組織的な犯罪に関与した容疑者が使う SNS のアカウントを AI を用いて解析し、人物相関図を作成して犯罪組織を可視化するというシステムを警察庁及び 5 都道府県警察で導入している<sup>14</sup>。

また官民連携の例として、2021 年、東日本旅客鉄道株式会社（JR 東日本）は、東京オリンピック・パラリンピック期間中に首都圏の一部駅において行動解析技術や顔認証技術を搭載した防犯カメラを活用し、不審者・不審物を検知するシステムを導入したが、テロなどの非常事態時には、カメラ画像が警察に非常時画像伝送システムを通じて伝送されるなどの連携も設計されていた<sup>15</sup>。

### 3.2 ツールキット展開に関する可能性と課題

日本の警察は、実証実験段階のものを含め、既に AI システムの導入に踏み切っているが、本ツールキットが提示する原則等は、警察における AI の使用に関する考え方と結論において

<sup>9</sup> 統合イノベーション戦略推進会議『AI 戦略 2022』（2022 年）p.28,

[https://www8.cao.go.jp/cstp/ai/aistrategy2022\\_honbun.pdf](https://www8.cao.go.jp/cstp/ai/aistrategy2022_honbun.pdf)

<sup>10</sup> 警察庁『令和 4 年版警察白書』（2022 年）p.12, <https://www.npa.go.jp/hakusyo/r04/index.html>

<sup>11</sup> 同庁においては、2022 年には同局を「サイバー警察局」に改組するなどの体制整備が行われている。

<sup>12</sup> 2022 年の組織改編により、同室は長官官房技術企画課の下に移管された。

<sup>13</sup> 斎藤正憲「警察における AI 等先端技術の導入に向けた取組について」『警察学論集』第 76 卷第 6 号（2023 年）

<sup>14</sup> 日本経済新聞電子版「SNS 解析システム導入へ AI 検索で人物相関図作成」（2021 年 5 月 30 日）、<https://www.nikkei.com/article/DGXZQOUE292FH0Z20C21A5000000/>

<sup>15</sup> 東日本旅客鉄道株式会社「東京 2020 オリンピック・パラリンピック競技大会に向けた鉄道セキュリティ向上の取り組みについて」。[https://www.jreast.co.jp/press/2021/20210706\\_ho02.pdf](https://www.jreast.co.jp/press/2021/20210706_ho02.pdf)

て一致する部分がある。例えば、本ツールキットが提示する原則のうち人間の自律性とは、人間に影響を与えるいかなる決定も、最終的には人間によって行われることを求めるものである。一方、先端技術導入企画室長を務めた齋藤正憲は、AI は一般に判定の過程や根拠を解明することが難しいものの、人間でなければできないとされてきた作業を一律的に行えるという特徴があることを挙げつつ、実証実験の結果を総括し、現段階で AI に 100% の精度を求ることは困難であり、人間による判断の支援を行うといった用途が適切である旨結論付けている<sup>16</sup>。

他方で、本ツールキットにおいて、リスク評価の方法やワークブックは、詳細かつ膨大な資料となっているが、どの組織がどの AI システムを使うかによって、必要な、又は重点的にみるべきリスクの評価方法は異なることから、本ツールキットをリスク評価に活用する際には、活用する主体、個別具体的な AI システムやサービス等を勘案して、参照可能な部分について検討を行う必要がある。例えば、日本警察では、AI システム等を使うユーザーとして警察庁と都道府県警察が想定されるが、個別の AI システム等の導入においては都道府県警察に裁量があり、その場合のリスク評価は都道府県警察が行うことが想定される。都道府県警察の中でも、AI システムの調達・開発に携わる部門、利活用する部門とで、必要とされるリスク評価の内容が異なると考えられる。また、都道府県警察の中でも、大規模県警察と小規模県警察とでは、AI システムの開発力、活用方法等に違いがある可能性がある。加えて、警察庁が開発した AI システムを都道府県警察が活用することも想定されるが、この場合、開発組織と利用組織が異なることから、組織体制に関するリスク評価が複雑となる。

このような観点から、本ツールキットを日本で参考としてリスク評価する際は、組織体制の影響評価といった構造的な評価ではなく、まず 2.2.3 で取り上げているような個別具体的な AI サービスに焦点をあて、それらが AI 特有のリスクを生じさせないかという具体的な観点での評価の参考として利用することも一案と考えられる。

今後、本ツールキットが提案する方法論や視点を日本の文脈や慣習に合わせた形で適用方法を検討することにより、日本の警察を含む法執行機関や、AI を利活用する公的機関や事業者が自らの組織における AI 導入を検討する際の参考となることを期待する。

---

<sup>16</sup> 齋藤正憲「警察における AI 等先端技術の導入に向けた取組について」『警察学論集』第 76 卷第 6 号（2023 年）