All Japan Educational Model United Nations

United Nations
General Assembly
1st Committee (DISEC)

EIGHTIETH UNITED NATIONS GENERAL ASSEMBLY FIRST COMMITTEE **A/80/1/DR.2**

Agenda item: AI と軍事 (AI and the military) **2025 年 8 月 5 日**

Sponsor: Argentina, Austria, Bangladesh, Brazil, Burkina Faso, Central African Republic, Chile, China, Costa Rica, Egypt, Eritrea, Estonia, Ethiopia, Fiji, France, Indonesia, Italy, Japan, Kenya, Kuwait, Latvia, Lebanon, Libya, Lithuania, New Zealand, Niger, Norway, Panama, Philippines, Poland, Slovakia, South Africa, Republic of Korea, Turkmenistan,

第 80 回国連総会第一委員会は、

2024 年 12 月 2 日に総会で採択された決議案 79/62 を想起し、

全ての兵器が国際人道法及び人権法にのっとるべきであるとする政府専門家会合(GGE)による LAWS に関して各国が考慮すべき指針を想起し、

国際司法裁判所(ICJ)の役割は、国際法に基づいて、国家間の法的紛争を裁くことだと認識し、

国際刑事裁判所(ICC)の役割は、戦争犯罪を裁くことだと認識し、

国際連合教育科学文化機構（UNESCO）のこれまでの AI に関する取り組みに好意を表明し、

特定通常兵器使用禁止制限条約(CCW)のこれまでの自律型致死兵器システムの規制に関する取り組みに好意を表明し、

LAWS の国際的に統一された定義が存在しておらず、それを定義することの必要性を認識し、

人間の介入なしに人命に関係する判断を自律的に下す AI 兵器は、国際人道法の原則に反し、決して用いられるべきでないことを確信し、

LAWS は自律性、致死力を持ち、人間の制御を超越する可能性のある前例のない兵器であると認識し、

LAWS は機械同士の争いではなく、人間の責任の帰属の問題も大きな争点になることに認識し、

LAWS の潜在的な危険性を憂慮し、

非政府組織による軍事用 AI の利用による被害を懸念し、

現状の国際枠組みが軍事 AI の誤作動や責任追及に十分対応できていない現状を考慮し、

非致死用途を含む平和目的の生成 AI の活用における国際協力を促し、

生成 AI 及び軍事 AI の急速な発展が国際社会、国際安全保障、人権、経済及び社会の安定に与える影響を認識し、

生成 AI 及び軍事 AI に関しての国際的な協力と法的拘束力を持つ枠組みの必要性を強調し、

AI リテラシーの普及は国ごとに差異があり、生成 AI の使用において AI リテラシーは必要不可欠であると認識し、

AI が学習するビッグデータに、知らないうちに個人情報が利用され、プライバシーを侵害するリスクがあることを認識し、

デジタルデバイドによって、国際安全保障にひずみが生じることに懸念を示し、

現状、生成 AI の開発に使用されているデータプールは一部の国によって開発され、情報に偏りがある可能性があることを考慮し、

AI が学習するデータによっては、特定の集団に対して差別的な傾向を含み、公平でないことを認識し、

AI による判断の根拠、仕組みが人間には理解できず、使用者、開発者が説明責任を果たせないことを認識し、

生成 AI が様々な要素および人間の関与によって形成されることを認識し、

生成 AI はそのリスクに応じた段階的な規制の分類が不可欠であると認識し、

現状、生成 AI 及び軍事 AI による被害、違法使用などの問題に対処する専門機関が存在しないことを憂慮し、

AI の多国家にわたる被害や違法使用に対処する新機関の必要性を認識し、

責任の所在を公平かつ中立的に判断する機関の必要性を認識し、

その判断には、AI に関する専門家や国際法学者といった多様な立場の人々の連携が不可欠であると認識し、

AI の運用において「意味のある人間の関与」が重要であることを確信し、

生成 AI を使用し作成したディープフェイクが名誉棄損及び誤情報の散布、プライバシーの侵害等の問題を引き起こすことを残念に思い、

1. LAWS とは目的の識別又は選定及び攻撃までを「意味ある人間の関与」なしに自律的に行える軍事用 AI であるとこの文書において定義し；
2. 国連事務総長に対し、期限である 2026 年末までに CCW で自律型致死兵器システムの規制に関する議定書が採択されたなかった場合、国連総会において条約交渉の場を設けることを要請する；
3. LAWS の定義について引き続き議論を続ける必要性を認識し；
4. 「意味のある人間の関与」を以下のように定義する：
    a. 自己破壊または自己無効化が可能であること,
    b. 攻撃時の影響が予測可能かつ攻撃が制御可能であること,
    c. 機能または攻撃任務がシステム単独で進化変更できないこと,
    d. 自律型致死兵器本体とその攻撃の影響が追跡可能で、説明可能であること,
    e. 攻撃の最終決定権が人間にあること；

5. 加盟国に対して、「意味のある人間の関与」を保証するために、軍事用 AI の使用において、以下の措置を講じるよう要請する：

    a. AI の作動プロセスにおいて、人間の統制が及ぶ段階と AI が自律的に動作する段階を明示する,

    b. AI が自律的に作動する段階以外では、人間の判断プロセスに影響を及ぼさないことを担保する；

6. 加盟国に対してこの決議案が採択される以前から、LAWS を保有していた場合は、国連事務総長に対して以下の内容の報告書を提出するよう強く促す：

    a. LAWS の保有数,

    b. LAWS の保有理由；

7. 加盟国に対し、法的拘束力のある規範または条約が採択されるまでの間、LAWS の使用及び開発に対するモラトリアムを実施するよう要請する；

8. 加盟国に対し、モラトリアム期間中に LAWS を使用した場合は、ICC において制裁及び賠償を議論する；

9. 加盟国に対し、LAWS が許可されている場合に誤作動が起きた場合は国家及び誤作動に至るまでの関係者が ICC にて裁かれることを確認する；

10. 加盟国に対し、ICC または ICJ で裁く基準を定めるために、新たな専門委員会を作ることを要請する；

11. 加盟国に対し、非政府組織が LAWS を保有・開発することを AI ハイレベル諮問機関において完全に禁止しすることを奨励し；

12. 加盟国に対し、世界の平和と安全を守るため、法的、技術的、倫理的、人道的、安全保障的観点の検討を含め、自律型兵器システム及び生成 AI がもたらすあらゆる課題と懸念に対処するため、包括的かつ包摂的なアプローチが必要となることを強調し；

13. 国連事務総長のもとおかれている「AI ハイレベル諮問機関」を以下の目的のために拡張する：

    a. 生成 AI における共同責任つまりは、事象ごとに責任の割合を変化させながら、開発者、提供者、利用者の共同で負うことに留意しつつ、どこにどれだけの賠償、制裁を行うかを決定する,

    b. AI 技術に関する研究を進め、その内容を国連総会の各委員会及び国際社会に公開する；

    c. AI の利用に関して監視を行う：

        (i) 生成 AI を用いた差別的、暴力的な表現を含む情報の拡散,

        (ii) 軍事用 AI の保有・使用について,

        (iii) AI による誤作動,

        (iv) 軍事 AI に関する技術移転を防ぐための監視；

e.(c)の研究データを参考にして、AI に関する会議を年に一回開催する：

(i)AI に関する規制を再検討,

(ii)AI が誤作動を起こした場合の対応策について再検討；

14. 加盟国に対し、UNESCO のグローバル AI 倫理・ガバナンス観測所のもとで以下の項目に関して多角的な視点のデータが取り入れられているグローバル AI データバンクを開発し、将来的には生成 AI 開発において最低限組み込まなければならないデータとして位置づけることを要請する：

    a. 地域固有の政治体制や戦争などの歴史的事実,

    b. 文化、言語、宗教に関する情報,

    c. 人権に関する法律および価値観；

15. 加盟国に対し、UNESCO の「グローバル AI 倫理とガバナンス観測所」に以下の要素を含む生成 AI に関する共通の倫理的基準を策定し、これを参考に各国で法整備することを要請する：

    a. プライバシーの尊重を最優先とし、ユーザーが自信のデータを学習に使用されない権利を持つこと及び顔や声など、個人を特定できる情報の模倣を制限すること,

    b. 差別的または攻撃的な発言を禁止し、フィルタリングを義務化すること,

    c. 兵器の製造、テロの計画、犯罪行為の手段など、法に反する行為を助長するコンテンツ生成を制限し、セキュリティリスクを防止すること,

    d. 医療、軍事、司法など人名や重大な人権にかかわる判断においては、AI の役割をあくまで補助的なものに限定し、最終判断は常に人間が行うべきであるということを呼びかけること,

    e. 印象操作な度に繋がるディープフェイクのフィルタリングを義務化し；

16. 加盟国に対し、AI と開発のためのグローバルパートナーシップのもとで、デジタルデバイド解消のための取り組みを進めることを奨励する；

17. 生成 AI に関して、以下の 5 段階に分類してリスクベースの規制を行うことを要請する：

    a. 禁止すべきリスク（道徳的・倫理的に問題があり、国際法上・倫理上許容できない AI 利用）：

        (i)バイアスを含み差別的な AI、違法な監視・強制的な行動操作を行う AI などが該当する,

        (ii)国際的に全面禁止とし、違反した際には制裁処置が取られる;

    b. 極高リスク（原則として、厳重な管理下でのみ使用が許される AI システム）：

        (i)国家規模の大規模監視 AI、原子力発電所・航空感性を完全自動化する AI などが該当する,

        (ii)開発前に、第三者機関による安全性審査・倫理審査を義務化する,

        (iii)年次報告書・外部監査を義務化する;

c. 高リスク（人間の生命・健康・権利に重大な影響を与える可能性があるが、リスク管理が可能な分野である AI）：

(i)医療、金融などの AI、重要インフラの制御 AI などが該当する,

(ii)生成 AI を使用して作成されたものに生成 AI マークの設定を義務付ける,

(iii)人間による最終判断を義務付ける;

d. 中リスク（社会への影響は限定的で、誤情報・偏見・依存リスクが存在するが、軽めの規制で対応可能である AI）：

i)画像・動画生成 AI などが該当する,

ii)生成 AI を使用して作成されたものに生成 AI マークの設定を義務付ける,

iii)出典を公開することを義務付ける;

18. 各国政府に対し、I の危険性についての教育を行うよう要請する.