

国連



地区制限あり

2024年3月11日

原語：英語

第78セッション

議題13

経済、社会および関連分野における国連の主要会議およびサミットの成果を統合的かつ協調的に実施し、フォローアップする。

アルバニア、アルゼンチン、オーストラリア、バハマ、ベルギー、ブラジル、ブルガリア、カボベルデ、カナダ、チリ、コートジボワール、クロアチア、チェコ、デンマーク、ジブチ、ドミニカ共和国、赤道ギニア、エストニア、フィジー、フィンランド、フランス、グルジア、ドイツ、ギリシャ、アイルランド、イスラエル、イタリア、日本、ヨルダン、ケニア、リベリア、ルクセンブルク、モルディブ、モンテネグロ、モロッコ、オランダ（王国）、ニュージーランド、北マケドニア、ペルー、大韓民国、ルーマニア、セルビア、シエラレオネ、シンガポール、スロバキア、スロベニア、南アフリカ、スペイン、スウェーデン、トルコ、アラブ首長国連邦、グレートブリテンおよび北アイルランド連合王国、米国、ウズベキスタン、ザンビア：\* 決議案

持続可能な開発のために、安全、安心、信頼できる人工知能システムのチャンスをつかむ

総会

国際法、特に国際連合憲章を再確認し、世界人権宣言を想起し、<sup>1</sup>

<sup>1</sup>決議 217 A (III)

「我々の世界を変革する：持続可能な開発のための 2030 アジェンダ」と題する 2015 年 9 月 25 日の決議 70/1、第 3 回開発資金国際会議のアジスアベバ行動アジェンダに関する 2015 年 7 月 27 日の決議 69/313、および 2023 年 9 月 29 日の決議 78/1 の附属書に含まれる、総会の主催の下で開催された持続可能な開発に関するハイレベル政治フォーラムで採択された政治宣言も再確認し、

持続可能な開発目標およびターゲットの達成に対する急速な技術変化の影響に関する 2023 年 7 月 25 日の決議 77/320、持続可能な開発のための情報通信技術に関する 2023 年 12 月 19 日の決議 78/132、持続可能な開発のための科学、技術およびイノベーションに関する 2023 年 12 月 19 日の決議 78/160、デジタル技術の文脈における人権の促進と保護に関する 2023 年 12 月 19 日の決議 78/213、デジタル時代におけるプライバシーへの権利に関する 2022 年 12 月 15 日の決議 77/211、情報社会に関する世界サミットの成果の実施の全体的なレビューに関する 2015 年 12 月 16 日の 70/125、ジュネーブ原則宣言、<sup>2</sup> ジュネーブ行動計画、<sup>3</sup> チュニス・コミットメント<sup>4</sup>、情報社会のためのチュニス・アジェンダ<sup>5</sup>、および 2020 年 9 月 21 日の決議 75/1 に含まれる国際連合創立 75 周年記念に関する宣言、を想起し、

国際電気通信連合が 40 の国連団体とパートナーシップを組み、年次サミットや、持続可能な開発目標を推進するための人工知能の責任ある実用的なアプリケーションを特定するための国際電気通信連合の人工知能リポジトリの立ち上げなど、「善のための人工知能 (Artificial Intelligence for Good)」プラットフォームを召集するための努力に留意する；国連教育科学文化機関総会による 2021 年 11 月 23 日の人工知能の倫理に関する勧告の採択、<sup>6</sup>、準備状況評価方法論および倫理的影響評価、人工知能の倫理に関するグローバル・フォーラムを含むその実施計画、ならびにビジネスと人権に関する指導原則に留意する；<sup>7</sup>、2011 年 6 月 16 日の人権理事会決議 17/4 で承認<sup>8</sup>、人工知能に関する国連人権高等弁務官事務所の作業にも留意し、

また、「デジタル協力のためのロードマップ」と題する国際連合事務総長報告書 (<sup>9</sup>)、ならびに、その実施を調整するための技術に関する事務総長特使事務所の設置、事務総長による人工知能に関するマルチ

---

\*スポンサーリストに変更があった場合は、公式記録に反映される。

<sup>2</sup>A/C.2/59/3、附属書 参照。

<sup>3</sup> 同上。

<sup>4</sup>A/60/687 を 参照のこと。

<sup>5</sup> 同上。

<sup>6</sup> 国際連合教育科学文化機関『総会記録』第 41 回総会、パリ、2021 年 11 月 9 日～24 日、第 1 巻、決議、附属書 VII。

<sup>7</sup>A/HRC/17/31、附属書。

<sup>8</sup>総会公式記録、第 66 会期別冊第 53 号 (A/66/53)、第 III 章を 参照のこと。III, sect.A.

<sup>9</sup>A/74/821.

ステークホルダーハイレベル諮問機関の設置、および 2023 年 12 月 21 日に発表されたその中間報告、その最終報告に留意し、

本決議では、非軍事領域の人工知能システムを指し、そのライフサイクルには以下の段階が含まれる：そのライフサイクルには、設計前、設計、開発、評価、試験、配備、使用、販売、調達、運用、廃棄の段階が含まれ、人間中心であり、信頼性があり、説明可能であり、倫理的であり、包括的であり、人権と国際法を十分に尊重し、推進し、保護するものであり、プライバシーを保護するものであり、持続可能な開発を志向するものであり、責任あるものである；また、人工知能システムの不適切または悪意ある設計、開発、配備、使用は、適切なセーフガードなし、または国際法と矛盾する方法など、持続可能な開発のための 2030 アジェンダとその持続可能な開発目標の達成に向けた進展を妨げ、経済、社会、環境の 3 つの側面における持続可能な開発を損ないかねないリスクをもたらすこと；情報の完全性と情報へのアクセスを弱体化させ、違法または恣意的なプライバシーの干渉を受けない権利を含む人権と基本的自由の保護、促進、享受を弱体化させ、事故の潜在的リスクと悪意ある行為者からの複合的脅威を増大させることを認識し、

さらに、人工知能システムの設計、開発、配備及び利用が急速に加速し、技術変化が急速に進んでいること、並びに、持続可能な開発目標の達成を加速する上で、人工知能システムが潜在的な影響を及ぼしていること、そのため、安全、セキュアかつ信頼できる人工知能システムに関するグローバルなコンセンサスを達成することの緊急性を強調し、イノベーションを促進し、安全、セキュアかつ信頼できる人工知能システムのガバナンスの分断を防ぐ、効果的で国際的に相互運用可能なセーフガード、慣行及び標準を策定し、利用するための包括的な国際協力を促進すること；また、現存する人工知能及びその他のデジタル・デバイド、並びに、各国間及び各国内における技術開発の様々なレベルを認識し、開発途上国がこの急速な加速に歩調を合わせる上で独自の課題に直面し、持続可能な開発の障害となっていること、条件、可能性及び能力の面で先進国と開発途上国の間に存在する格差を縮小する必要性を認識し、従って、各国間及び各国内のデジタル・デバイドを解消するため、開発途上国に対する能力構築及び技術的・資金的支援を強化し、人工知能システムのガバナンスに関する国際プロセス及びフォーラムへの開発途上国の効果的、公平かつ有意義な参加及び代表を支援することの緊急性をも強調することを認識し、

また、人工知能システムのガバナンスは発展途上の分野であり、技術とそれに対する我々の理解が発展するにつれて、国際法に基づき、相互運用可能で、機敏で、適応可能で、包括的で、先進国・途上国を問わず様々なニーズと能力に対応し、全ての人の利益のために、適切なガバナンス・アプローチの可能性について継続的に議論する必要があることをガバナンスは認識し、

1. 人工知能やその他のデジタル分野における国家間・国内間の溝を埋めることを決議する；

2. 持続可能な開発のための 2030 アジェンダの完全な実現に向けた進展を加速させるため、安全でセキュアで信頼できる人工知能システムを推進すること<sup>10</sup>、各国間および各国内の人工知能およびその他のデジタル・デバイドをさらに埋める。また、17 の持続可能な開発目標すべてと、経済、社会、環境の 3 つの側面における持続可能な開発を達成し、特に途上国にとって共有されるその他のグローバルな課題に対処するため、デジタルトランスフォーメーションとその恩恵への公平なアクセスを妨げるのではなく、促進する安全でセキュアで信頼できる人工知能システムの標準の必要性を強調することを決議する；
3. 加盟国に対し、民間セクター、国際機関、地域機関、市民社会、メディア、学術研究機関、技術コミュニティ、個人を含む、それぞれの役割と責任の範囲内で、すべての地域と国のマルチステークホルダーに対し、安全な人工知能システムに関する規制とガバナンスのアプローチと枠組みを開発し、支援するよう奨励し、要請する、このようなアプローチ及び枠組みを策定するにあたっては、ガバナンスとマルチ・ステークホルダー間の効果的なパートナーシップ及び協力が必要であることを認識しつつ、イノベーション、起業家精神、及び相互に合意された条件による知識・技術の普及を含め、あらゆるレベルにおいて実現可能なエコシステムを構築する、安全で信頼できる人工知能システムに関連する規制及びガバナンスのアプローチ及び枠組みを策定し、支援する；
4. 加盟国に対し、デジタル変革の恩恵と安全、安心かつ信頼できる人工知能システムへの包括的かつ公平なアクセスに向けて、開発途上国と協力し、援助を提供するための行動をとるよう呼びかけ、他の利害関係者に対し、以下を含む行動をとるよう求める：
  - (a) すべての国、特に開発途上国のデジタル・トランスフォーメーションへの参加を拡大し、恩恵を活用し、安全、安心かつ信頼できる人工知能システムの開発、配備、利用に効果的に参加する。これには、人工知能システムに関する能力構築も含まれ、知識共有活動や相互に合意した条件での技術移転を促進することが能力構築の重要な側面であることを認識し、人工知能やその他のデジタル・デバイドを解消する必要性を強調する；
  - (b) 開発途上国が人工知能システムのライフサイクルを通じて効果的に参加し、2030 アジェンダとその持続可能な開発目標の完全な実現に向けてなど、人工知能システムの社会への包括的かつ積極的な貢献を加速できるよう、パートナーシップの強化を通じて、デジタル・インフラの接続性と技術革新へのアクセスを強化する；

---

<sup>10</sup>決議 70/1

- (c) 開発途上国、特に後発開発途上国が、17 の持続可能な開発目標すべてを達成するために、主要な構造的障害に対処し、科学的情報源、手頃な価格の技術、研究開発の利用を拡大することを通じ、パートナーシップの強化を含め、新技術および新興技術、人工知能イノベーションの恩恵にアクセスするための障害を取り除く能力を強化する；
- (d) 持続可能な開発目標に関連し、デジタル技術と安全・安心・信頼できる人工知能システムに関連する研究とイノベーションのための資金を増やし、この研究に貢献し、そこから利益を得るための能力をすべての地域と国で構築することを目指す；
- (e) 技術的専門知識と能力を開発し、データと計算資源を活用し、国の規制とガバナンスのアプローチ、フレームワーク、調達能力を開発する途上国の能力を強化し、安全、安心、信頼できる人工知能システムをベースとしたソリューションのための包括的な実現環境をあらゆるレベルで構築するための、国際的なイノベーションに基づく環境を可能にする；
- (f) 開発途上国の国家的ニーズ、政策、優先事項に従って、相互に合意された条件での技術移転、人工知能やその他のデジタル・デバイスを開発するための能力構築、技術支援、人工知能システムに関する開発途上国への融資など、実施手段を早急に動員する；
- (g) 経済、社会、環境の3つの側面における持続可能な発展を達成するために、安全、安心、信頼できる人工知能システムへのアクセス、設計、開発、配備、利用を促進する；
5. 人工知能システムのライフサイクル全体を通じて、人権と基本的自由が尊重され、保護され、促進されなければならないことを強調し、国際人権法を遵守して運用することが不可能な、あるいは人権、特に脆弱な状況にある人々の人権の享受に過度のリスクをもたらすような人工知能システムの使用を控える、あるいは中止するよう、すべての加盟国および適用法がある場合にはその他の利害関係者に呼びかけ、人工知能システムのライフサイクル全体を含め、人々がオフラインで有する権利と同じ権利がオンラインでも保護されなければならないことを再確認する；
6. すべての加盟国に対し、適切な場合には、それぞれの国の優先事項や状況に沿い、それぞれの国の規制やガバナンスのアプローチや枠組みを実施しながら、また、該当する場合には、他の利害関係者に対し、包括的かつ衡平な方法で、万人の利益のために、安全、安心、信頼できる人工知能システムを推進し、発展途上国に特別な配慮をした、経済、社会、環境の3つの側面における持続可能な開発の達成や、誰一人取り残さないことを含む、世界の最大の課題に取り組むための、そのようなシステムを可能にする環境を醸成することを奨励する；
- (a) 持続可能な開発のための責任ある包括的な人工知能のイノベーションと投資を支援すると同時に、安全、安心かつ信頼できる人工知能システムを促進するため、それぞれの国、および適用法があ

る場合にはその国の政策と優先事項、ならびに国際法上の義務に沿った、国内の規制およびガバナンスのアプローチと枠組みの開発と実施を促進する；

- (b) 人工知能システムの設計・開発中および配備・使用前に、脆弱性とリスクを国際的に相互運用可能な形で特定、分類、評価、テスト、予防、軽減するための技術革新を促進する効果的な対策を奨励する；
- (c) 人工知能システムおよび人工知能インシデントの開発、テスト、配備後の技術的脆弱性、および必要に応じて誤用について、エンドユーザーおよびサードパーティによる証拠に基づく発見と報告を可能にするフィードバック・メカニズムを組み込むことを奨励する；
- (d) 人工知能システムの適切な市民利用の中核となる機能、能力、限界、領域について、一般市民の意識と理解を高める；
- (e) 人工知能システムのライフサイクル全体にわたって、リスクモニタリングと管理の仕組み、個人データ保護とプライバシーポリシーを含むデータセキュリティの仕組み、および適切な影響アセスメントの開発、実施、開示を促進する；
- (f) 物理的セキュリティ、人工知能システムのセキュリティ、人工知能システムのライフサイクル全般にわたるリスクマネジメントを含む、効果的なセーフガードの開発と実施への投資を強化する；
- (g) 利用者が情報操作を識別し、真正なデジタルコンテンツと生成的人工知能によって生成または操作されたデジタルコンテンツの出所を区別または特定することを可能にする、信頼できるコンテンツ認証および出所証明メカニズム（技術的に実現可能かつ適切な場合には、電子透かしまたはラベリングなど）を含む、効果的で、利用しやすく、適応可能で、国際的に相互運用可能な技術的ツール、標準または慣行の開発および展開を奨励し、メディアおよび情報リテラシーを向上させる；
- (h) 政策立案を強化し、あらゆる形態の差別、偏見、誤用、その他の危害から個人を保護するために、人工知能システムの訓練とテストのための効果的で国際的に相互運用可能な枠組み、慣行、標準の開発と実施を促進する、人工知能システムのライフサイクル全体を通じて、差別的または偏ったアプリケーションと結果を強化または永続化することを回避すること。例えば、データセットに符号化されたバイアスを分析・低減し、アルゴリズムによる差別や偏見と闘う一方で、他のユーザーや受益者の積極的な開発、アクセス、利用に不注意または不釣り合いな影響を与えないようにすること；

- 
- (i) 適切かつ適切な場合には、イノベーションを促進しつつ、著作権で保護されたコンテンツを含む知的財産権を尊重するための適切なセーフガードの実施を奨励する；
  - (j) 人工知能システムのライフサイクル全体を通しての個人データの利用を含め、システムのテストや評価を行う際のプライバシー・個人データの保護、および適用される国際的、国内的、サブナショナルな法的枠組みを遵守した透明性と報告要件の保護；
  - (k) エンドユーザーに影響を与える決定を行う、またはそれを支援する人工知能システムのライフサイクル全体を通じて、透明性、予測可能性、信頼性、理解可能性を促進する。これには、通知と説明を提供すること、自動化された決定と関連プロセスのレビューや、適切かつ関連性がある場合には、人間による意思決定の代替、または人工知能システムの自動化された決定によって悪影響を受けた人々に対する効果的な救済と説明責任など、人間による監視を促進することが含まれる；
  - (l) 人権と基本的自由の完全かつ効果的な享有を保護し、潜在的な影響を緩和するために、人工知能システムのライフサイクル全体を通じて、リスクと影響のアセスメントを含む効果的なセーフガードの開発と実施への投資を強化する；
  - (m) 言語的・文化的多様性を促進・保護・保全する人工知能システムを防御し、特に大規模な言語モデルについては、学習データや人工知能システムのライフサイクル全体を通して多言語主義を考慮する；
  - (n) 人工知能システムのライフサイクル全体で役割を担う事業者間で、相互に合意した条件で情報共有を低減し、科学的かつ証拠に基づく人工知能システムのベストプラクティス、方針、アプローチを特定、理解、活用して行動することで、高度な人工知能システムを含む人工知能システムのライフサイクル全体で、潜在的リスクを軽減しながら利益を最大化する；
  - (o) オープンソースの人工知能システムなどのデジタル・ソリューションのスケールアップの役割を含め、デジタル・デバイドを解消し、17 の持続可能な開発目標すべてを達成する上での人工知能システムの役割に関連する潜在的な利益とリスクを理解し、バランスをとり、対処するための研究と国際協力を奨励する；
  - (p) 加盟国に対し、ジェンダー・デジタル・デバイドを解消するための具体的な措置を採用し、アクセス、購入可能な価格、デジタル・リテラシー、プライバシー、オンラインの安全性に特に注意を払い、人工知能システムを含むデジタル技術の利用を強化し、政策決定やそれを導く枠組みにおいて障害、ジェンダー、人種平等の視点を主流にするよう求める；

- (q) 研究および国際協力を奨励し、人工知能システムの導入が労働市場に及ぼす影響を特定・評価するための手段を開発し、特に開発途上国、特に後発開発途上国において、労働力に対する潜在的な悪影響を軽減するための支援を提供し、デジタル研修、能力開発、技術革新の支援、人工知能システムの恩恵へのアクセス強化を目的としたプログラムを育成する；
7. データは人工知能システムの開発と運用の基本であることも認識する；加盟国に対し、データ・ガバナンスに関するベスト・プラクティスを共有し、安全でセキュアで信頼できる人工知能システムのための信頼できる国境を越えたデータの流れを促進するためのアプローチの一貫性と相互運用性を高めるため、実行可能な場合には、データ・ガバナンスに関する国際的な協力、協調、支援を促進し、その開発をより包括的、公平、効果的かつ万人にとって有益なものにするよう求める；
  8. 国際的なアプローチが人工知能システムとその用途の進化に歩調を合わせられるよう、人工知能ガバナンスの分野における発展に関する議論を継続することの重要性を認識する；人工知能システム及び急速な技術革新が、既存及び新たな技術の開発において、また、17の持続可能な開発目標全ての達成を加速する上でもたらし得る潜在的な影響と応用について、全ての関係者に利益をもたらす包括的な研究、マッピング及び分析を推進し、安全、安心かつ信頼できる人工知能システムのために、人工知能の設計者、開発者、評価者、配備者、利用者及びその他の利害関係者のために、効果的で国際的に相互運用可能な保障措置、慣行、標準及びツールを開発、促進及び実施する方法を知らせるために、国際社会が継続的に努力することを奨励する；また、政府、民間セクター、市民社会、国際・地域機構、学術・研究機関、技術コミュニティ、その他すべての利害関係者が、必要に応じて引き続き協力する必要性を強調するとともに、安全、セキュアかつ信頼できる人工知能システムの包括的ガバナンスにおいて、すべてのコミュニティ、特に途上国のコミュニティが、より結束し、効果的、協調的かつ包括的に関与・参加する必要性を認識する；
  9. 適用される国際法および国内法を遵守し、国連「ビジネスと人権に関する指導原則」に沿って行動するよう民間セクターを奨励する；また、安全でセキュアで信頼できる人工知能システムの恩恵を、より包括的かつ公平に享受することの重要性を認識し、安全でセキュアで信頼できる人工知能のライフサイクル全体にわたって、公正でオープン、包括的かつ差別のないビジネス環境、経済・商業活動、競争力のあるエコシステムおよび市場を提供・促進するために、官民および市民社会、学術・研究機関、技術コミュニティの間やその内部を含め、協力を強化する必要性を認識する；また、加盟国に対し、特に発展途上国に対する重要な投資を通じて、中小企業や起業家、技術的才能を持つ人々に対する新たな機会を支援し可能にすること、および人工知能市場における公正な競争を可能にすることを含め、安全、安心、かつ信頼できる人工知能システムおよび関連技術における競争を促進するための政策および規制を策定するよう奨励する；



10. 国際連合システムの専門機関、基金、計画、その他の事業体、団体および事務所、ならびに関連機関に対し、それぞれの権限と資源の範囲内で、潜在的な影響と応用について全ての関係者に有益な調査、マッピングおよび分析を行うことを含め、適切な機関間メカニズムを通じ、協力的、協調的かつ包括的な方法で、人工知能システムがもたらす機会を活用し、課題に対処するための対応を引き続き評価し、強化するよう求める；また、17の持続可能な開発目標すべてと、経済、社会、環境の3つの側面における持続可能な開発の達成に向け、能力開発、安全で信頼できる人工知能システムへのアクセス、およびその恩恵の共有において、途上国と協力し、途上国を支援する；
11. 2022年9月8日の決議 [76/307](#) および 2023年9月1日の決定書 77/568 を想起し、未来のサミットの様式と範囲に関し、世界的なデジタル・コンパクトの策定に期待する；
12. 2025年、情報社会に関する世界サミット以降の進展について、総会による全体的な見直しが行われることも期待している；
13. 国際連合システムは、そのマンデートに則り、国際法、特に国際連合憲章、世界人権宣言、および持続可能な開発のための2030アジェンダに合致した、安全、安心かつ信頼できる人工知能システムに関する世界的な合意形成に独自に貢献する。