

国内外の動向及び国際的な議論の動向

令和2年11月12日
事 務 局

国内の動向

AI戦略 2019（2020.06フォローアップ版）【2020年6月22日】

- AI戦略 2019（2020.06フォローアップ版）において、倫理について、次のように記載。
 - ・ 責任あるAIやイノベーション等の推進に向けた、GPAIの設立と専門家派遣を通じた議論への貢献（2020年度）【総・経】
 - ・ AI社会原則の実装に向けて、国内外の動向も見据えつつ、我が国の産業競争力の強化と、AIの社会受容の向上に資する規制、標準化、ガイドライン、監査等、我が国のAIガバナンスの在り方を検討（2020年度）【CSTI・総・経】 等

機械学習品質マネジメントガイドライン【2020年6月30日】

- 新エネルギー・産業技術総合開発機構（NEDO）は、2020年6月30日に、機械学習を用いたAIシステムの品質に関して、客観的に評価できる「機械学習品質マネジメントガイドライン」の第一版を策定し、公表。
 - ・ 内部品質（機械学習要素が固有に持つ特性）の向上により外部品質（AIシステム中で機械学習要素に要求される品質）が必要なレベルで達成され、最終的な製品における利用時品質（AIシステム利用時に必要な品質）が実現される。
 - ・ 外部品質として、(1)リスク回避性、(2)AIパフォーマンス、(3)公平性の3つの軸を設定。内部品質として、【1】要求分析の十分性、【2】データ設計の十分性、【3】データセットの被覆性、【4】データセットの均一性、【5】機械学習モデルの正確性、【6】機械学習モデルの安定性、【7】プログラムの健全性、【8】運用時品質の維持性の8つから、(1)～(3)の外部品質が充足しているか判断。

世界最先端デジタル国家創造宣言・官民データ活用推進基本計画【2020年7月17日】

- 2020年7月17日付けで閣議決定された「世界最先端デジタル国家創造宣言・官民データ活用推進基本計画」において、AIネットワーク化について、次のとおり記載
 - ・ 「AIネットワーク社会推進会議」において、G7、G20、OECD等の国際的な議論の動向を踏まえつつ、AIの社会実装の推進により、どのような社会経済を目指すべきか、基本的な政策や中長期的な戦略の在り方について検討する。
 - ・ G7、G20、OECD等がAIの社会実装に向けた様々な検討を現在進めているところ、日本の検討状況や成果を提言すること等を通じてグローバルなAI社会の実現に貢献する。 等

海外の動向

欧州委員会 AI白書に関する意見募集の結果を公表

- 欧州委員会は、2020年2月19日に公表したAI白書について、同年5月19日まで意見募集を実施（6月14日まで延長）。
- 意見募集の結果については、同年7月17日に公表。世界中から1,215の意見が提出され、規制オプションに関する主な関心事項としては、次のとおり概要をまとめている。
 - ・ 提出された意見のうち、42%が新たな規制枠組みの導入を求めている。
 - ・ 提出された意見のうち、43%が義務的な要求はハイリスクAIに限るべきと考えている。
- 意見募集や開始影響評価※の結果を踏まえ、2021年の早い時期に規制枠組みの提案がなされる予定。
 - ※ 規制の提案に向けた政策評価プロセスの一環であり、意見募集を受けて、EUは政策オプションごとに経済等への影響を審査することとなっている。

欧州委員会 「信頼できる人工知能の評価リスト（ALTAI）」を公表

- 欧州委員会により選定されたAIハイレベル専門家グループ（High-Level Expert Group on Artificial Intelligence（AI HLEG））は、2019年4月に取りまとめた非拘束的な「信頼できるAIのための倫理ガイドライン」の一部として、自己評価のための評価リストを策定。

【倫理ガイドラインの4つの倫理原則】

- ① 人間の自立性の尊重
- ② 危害の防止
- ③ 公平性
- ④ 説明可能性

【倫理ガイドラインの7つの要求条件】

- ① 人間の活動と監視
- ② 堅固性と安全性
- ③ プライバシーとデータガバナンス
- ④ 透明性
- ⑤ 多様性、非差別、公平性
- ⑥ 社会・環境の幸福
- ⑦ アカウンタビリティ

- AI HLEGは、2019年後半に評価リストを試行した団体からのフィードバックを受け、この評価リストを見直し、2020年7月に最終評価リストを公表。

海外の動向

欧州議会 「人工知能、ロボットおよび関連技術の倫理的側面の枠組み」を採択

- 欧州議会は、2020年10月20日に、「人工知能、ロボットおよび関連技術の倫理的側面の枠組み」を採択。
- 将来の法律で考慮されるべき指針に関する提案であり、ハイリスクAI技術については、随時、人間の監視下にあるようにデザインされる必要があるとしている。

【主な指針項目】

人間中心のAI、リスクアセスメント、安全機能・透明性及び説明責任、非偏見及び非差別、社会的責任とジェンダーバランス、環境と持続可能性、プライバシーと生体認証、良いガバナンス、セキュリティと防御、輸送、雇用・労働者の権利・デジタルスキル・職場、教育と文化、国家監督当局、連合レベルでの調整、倫理的コンプライアンスのヨーロッパ認証、国際協力

欧州議会 「人工知能の民事責任レジーム」を採択

- 欧州議会は、2020年10月20日に、「人工知能の民事責任レジーム」を採択。
- 未来志向の民事責任枠組みのための指針に関する提案であり、ハイリスクAIのオペレーターに対して、損害に関する厳格な責任を適用すること等が記載されている。明確な法的枠組みは、市民を守り、AI技術への信頼を促進する一方、ビジネスに法的な確実性をもたらし、イノベーションを促進するとしている。

【主な指針項目】

責任とAI、オペレーターの責任、異なるリスクに応じた異なる責任ルール、保険とAIシステム

(注) 欧州議会では、上記2件に、知的財産権に関する報告書を加えた合計3件のAI関連の提案を同時に採択した。

海外の動向

英国 情報コミッショナーズオフィス(ICO)「AIとデータ保護に関するガイダンス」を公表

- 英国 情報コミッショナーズオフィス (ICO) は、2020年7月30日に、「AIとデータ保護に関するガイダンス」を公表。
- AIにおけるデータ保護の実践を実現する取組の一環として取りまとめられたもので、AIによって悪化あるいは新たに生じるセキュリティリスク、差別やバイアスなど様々なリスクを軽減するために、組織が活用できるベストプラクティスと技術的対策に関する推奨事項が掲載。

【ガイダンスの構成】

- ・ データ保護の影響評価を含む、AIにおける説明責任とガバナンス
- ・ 合法的な根拠、AIシステムのパフォーマンスの評価と改善、潜在的な差別の緩和など公正で合法的かつ透明性のある処理
- ・ データの最小化とセキュリティ
- ・ 自動化された意思決定に関連する権利を含む個人の権利の遵守

米国 国立標準技術研究所 (NIST) 説明可能な人工知能の4原則のドラフトを公表

- 米国 国立標準技術研究所 (NIST) は、2020年8月18日に、説明可能な人工知能の4原則のドラフトを公表し、10月15日まで意見募集を実施。

【4原則】

- ・ Explanation (説明) : AIシステムは、全ての出力に付随する証拠や理由を提供する必要がある。
- ・ Meaningful (有意味) : システムは、個々の利用者が利用できる説明を提供する必要がある。
- ・ Explanation Accuracy (説明の正確性) : 説明は、出力を生成するためのシステムプロセスを正しく反映している。
- ・ Knowledge Limits (知識の限界) : システムは、それが設計された条件下でのみ又はシステムがその出力が十分に信頼できる状況になった場合にのみ動作する。

国際的な議論の動向

インターネットエコノミーに関する日米政策協力対話【2020年9月17日、18日】

- AIセッションにおいて、信頼のおけるAIの責任あるガバナンスを支援する官民の議論を実施。総務省からは「報告書2020」の概要資料を共有。
 - ・ 日本及び米国は、世界貿易機関（WTO）国際電気通信連合（ITU）、経済協力開発機構（OECD）、アジア太平洋経済協力（APEC）、G20、G7、インターネットガバナンスフォーラム（IGF）等の場におけるものを含め、グローバルなデジタル経済政策環境を向上するための国際場裡における緊密な協力についての継続的なコミットメントを強調。
 - ・ 両国は、マルチステークホルダー・アプローチに基づいた、包括的で、開かれた、透明性のあるインターネットガバナンス・システムへのコミットメントを再確認。また、両国は、OECDのAIに関する勧告への支持を再確認し、継続的にOECDやGPAIを通じてAIについて協働することを確認。

日EU ICT戦略ワークショップ【2020年10月1日、2日】

- AIセッションにおいて、双方の政策動向やビジネス現場での課題に関する意見交換を実施。総務省からは「報告書2020」の概要を説明。
 - ・ EUが2020年2月に公表したAI白書で示されているAIの定義やハイリスクAIの厳密な定義が必要であることが日EU双方から指摘され、引き続き、様々なステークホルダーで議論を深めていくことが求められた。

OECD（デジタル経済政策委員会：CDEP）

- 2020年11月末に、CDEP AIセッションを開催予定。
- ONE AI（AI専門家会合）傘下のWG（Classification of AI systems, Implementing trustworthy AI, National AI policies）における議論について、中間レポートを含む進捗報告を予定。その他、OECD.AI（AI政策に関するオブザーバトリ）についてのアップデート及びデモンストレーション等がなされる見込み。

国際的な議論の動向

GPAI (Global Partnership on AI)

○ 設立趣旨

人権の尊重・包摂性・多様性・イノベーション・経済成長を主眼とした「責任あるAI」を政府・国際機関・産業界・専門家等からなるマルチステークホルダー方式で推進する。

○ 設立経緯

2019年ビアリッツサミット（フランス）において、GPAIの立ち上げが提唱。2020年5月のG7科学技術大臣会合において、立ち上げに関するG7の協力に合意。その後、以下の参加国によって同年6月15日に立ち上がった。

○ 参加国

G7メンバー国・機関全て。その他、オーストラリア、ニュージーランド、スロベニア、韓国、シンガポール、メキシコ、インド。

○ 取組内容

2020年7月以降、以下の5つのテーマについてWGを設置し、12月のプレナリー会合に向けて検討中。

- ・ 責任あるA I
- ・ データガバナンス
- ・ 仕事の未来
- ・ イノベーションと商業化
- ・ 新型コロナウイルス感染症への対応（※） （※） 「責任あるA I」WGのサブグループとして設置されている。

ユネスコ (UNESCO)

- 2021年末の第41回総会でのAI倫理に関するグローバル勧告の策定を目指し、2020年3月に設置した専門家グループ（AHEG）の第2回会合を開催（8月31日～9月4日）。第1次ドラフトについて、2020年末まで各国にコメント照会中、2021年からは政府間会合が開催予定。

国際的な議論の動向

欧州評議会 AIに関するアドホック会合（CAHAI）

- 議長国フランスのイニシアチブの下で、2019年9月の閣僚代理会合で立ち上げを採択。
- 各加盟国において、AIの開発・利用に際して人権・民主主義の保護が十分に確保されているかを審査するための法的枠組の制定の検討を主目的とする
- 2020年9月、これまでの進捗（各国・各機関における法的枠組み等のマッピング、テーマ別WGの設置等）をまとめた第1次プロGRESSレポートが、閣僚代理会合にて採択された。

Partnership on AI

- Partnership on AIは、2020年7月30日に、「人工知能と感情知性に関する倫理」と題するレポートを公表。
- 本レポートでは、感情等を対象としたAIや、それに伴って生じる倫理的課題について、産業界、学术界、政府、市民社会等の様々な分野との対話に基づきまとめている。

【主なトピック】

感情コンピューティングの定義及び分類、感情コンピューティングで用いられる人間の感情・センサー・データの種類 等

国連 自動車基準調和世界フォーラム（WP29）

- 国連 自動車基準調和世界フォーラム（WP29）において、2020年6月24日に、自動運行装置（レベル3）に係る国際基準が初めて成立。

【自動運行装置の国際基準の主な要件】

- ・ 少なくとも注意深く有能な運転者と同等以上のレベルの事故回避性能
- ・ ドライバーモニタリングシステムの搭載。システムの作動状態記録装置の搭載 等

（注）日本国内においては、今回の国際基準の成立に先行して、2020年4月1日に、同等の内容の自動運行装置等の基準を施行済。



EU–Japan Digital Strategies Workshop

High-Level Expert Group and White Paper on AI

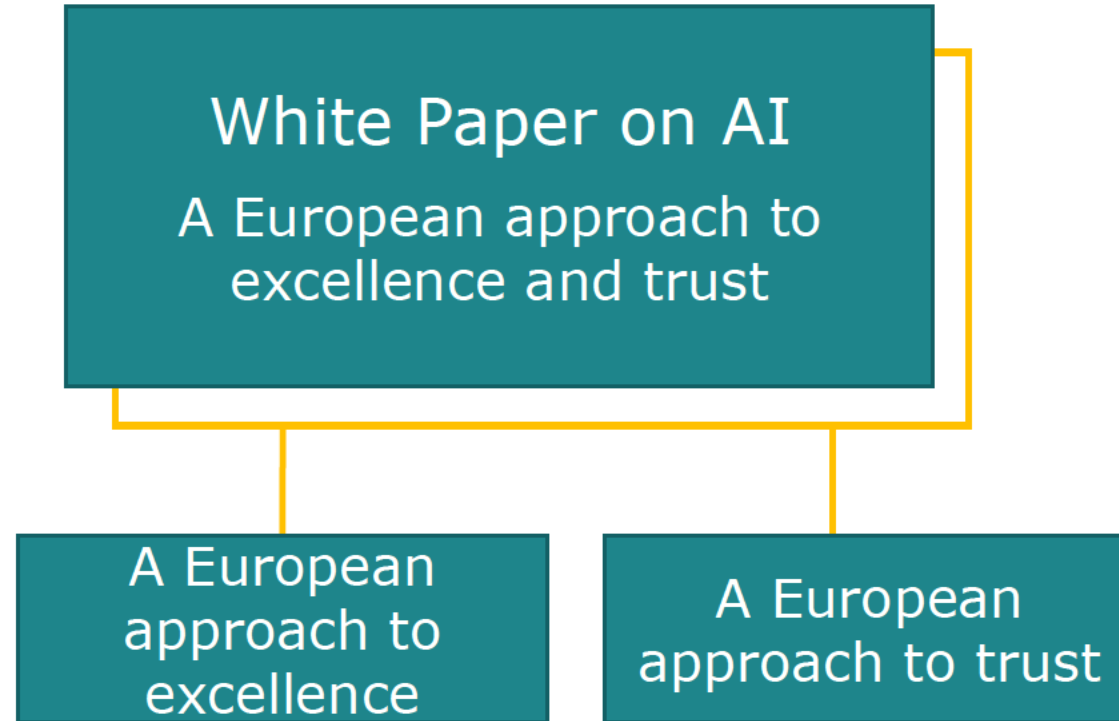
1 October 2020

Juha Heikkilä, PhD
Head of Unit
Robotics & Artificial Intelligence

High-Level Expert Group on AI

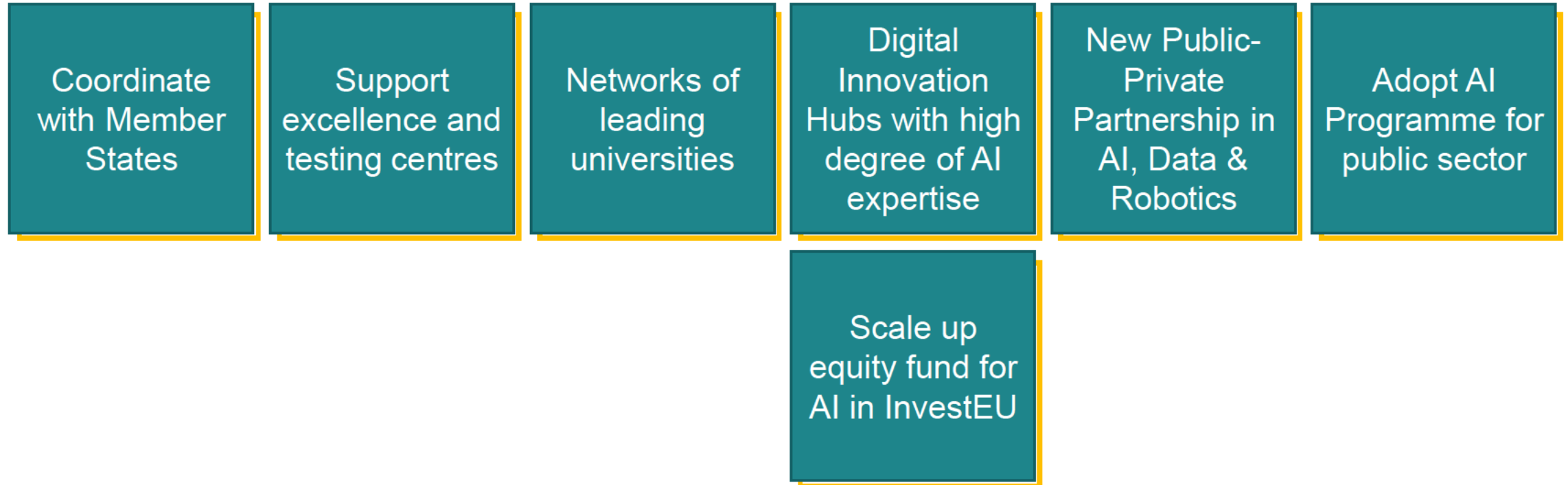
- April 2019: Ethics Guidelines for Trustworthy AI
 - 7 key ethics requirements
 - Draft Assessment List to put requirements into practice
- June 2019: Policy and Investment Recommendations
- July 2020: final deliverables
 - Final Assessment List for Trustworthy Artificial Intelligence, also as a web-tool
 - Sectorial Considerations on Policy and Investment Recommendations for Trustworthy AI

February 2020: a balanced approach to AI



Foster an Ecosystem of Excellence

Leverage EU strengths to accelerate AI development, application and use




The ecosystem of excellence builds on the Coordinated Plan on AI (2018).

Promote an Ecosystem of Trust

Risk-based and proportionate regulatory approach could include

- Identification of **high-risk AI systems**
- **Mandatory requirements** and **ex-ante conformity assessment** for high-risk AI systems
- For AI systems that are not high-risk, a **voluntary labelling scheme**

Realising an ecosystem of trust prevents fragmentation of the single market



Possible requirements:

- ✓ Training data
- ✓ Data and record keeping
- ✓ Information to be provided
- ✓ Robustness and accuracy
- ✓ Human oversight

Public consultation on the AI White Paper attracted broad interest

Responses received
February –
June 2020

1,215

84% of responses were from **EU**

- Responses also from India, China, Japan, Syria, Iraq, Brazil, Mexico, Canada, the US and the UK



Public consultation: ecosystem of trust

Need for new legislation

- **Only 3%** think that **current legislation is fully sufficient.**
- **33%** find that the **current legislation may have some gaps.**
- **42%** say there is a **need for a new legislation**

Focused on high-risk?

- **43%** of respondents agree that **compulsory requirements should be limited to high-risk AI applications.**
- **31%** do not agree with the limitation to high-risk

Main concerns of respondents:

90%

Breaching fundamental rights

87%

Discriminatory outcomes

Next steps

We are **analysing** the public consultation results and preparing an **impact assessment**

A **revised Coordinated Plan on AI** and a proposal for a **regulatory framework** could be put forward early in 2021

AI Alliance Assembly
on 9 October for
stakeholders

**Inter-Service
consultation**