

# AIに関する国際的議論

中央大学 実積寿也



# 世界が注目する日本の議論

## 国際的な議論のための AI開発ガイドライン案 (総務省、2017)

1. 連携の原則
2. 透明性の原則
3. 制御可能性の原則
4. 安全性の原則
5. セキュリティの原則
6. プライバシーの原則
7. 倫理の原則
8. 利用者支援の原則
9. アカウンタビリティの原則

## AI利活用原則案 (総務省、2017)

1. 適正利用の原則
2. 適正学習の原則
3. 連携の原則
4. 安全の原則
5. セキュリティの原則
6. プライバシーの原則
7. 尊厳・自律の原則
8. 公平性の原則
9. 透明性の原則
10. アカウンタビリティの原則

OECD  
/CDEP

## 人工知能に関する理事会勧告 (OECD, 2019)

信頼できるAIの責任あるスチュワードシップのための原則

1. 包摂的な成長、持続可能な開発及び幸福
2. 人間中心の価値観及び公平性
3. 透明性及び説明可能性
4. 頑健性、セキュリティ及び安全性
5. アカウンタビリティ

信頼できるAIのための国内政策と国際協力

1. AIの研究開発への投資
2. AIのためのデジタル・エコシステムの整備
3. AIを推進するための政策環境の形成
4. 人材育成及び労働市場の変化への備え
5. 信頼できるAIのための国際協力

## 人工知能学会倫理指針(人工知能学 会、2017)

1. 人類への貢献
2. 法規制の遵守
3. 他者のプライバシーの尊重
4. 公正性
5. 安全性
6. 誠実な振る舞い
7. 社会に対する責任
8. 社会との対話と自己研鑽
9. 人工知能への倫理遵守の要請

## 人間中心のAI社会原則(統合イノベー ション戦略推進会議、2019)

AI社会原則

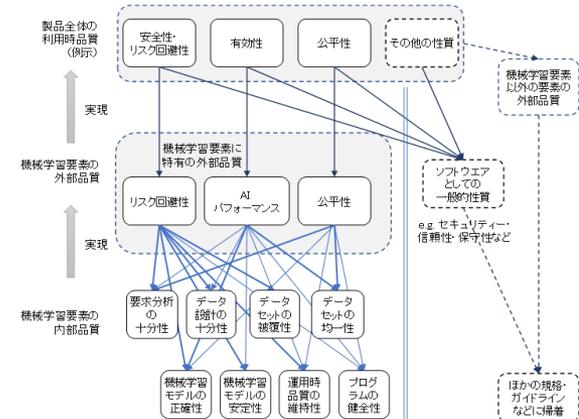
1. 人間中心の原則
2. 教育・リテラシーの原則
3. プライバシー確保の原則
4. セキュリティ確保の原則
5. 公正競争確保の原則
6. 公平性、説明責任及び透明性の原則
7. イノベーションの原則

AI開発利用原則

早急にオープンな議論を通じて国際的なコンセンサスを醸成し、非規制的で非拘束的な枠組みとして国際的に共有されることが重要

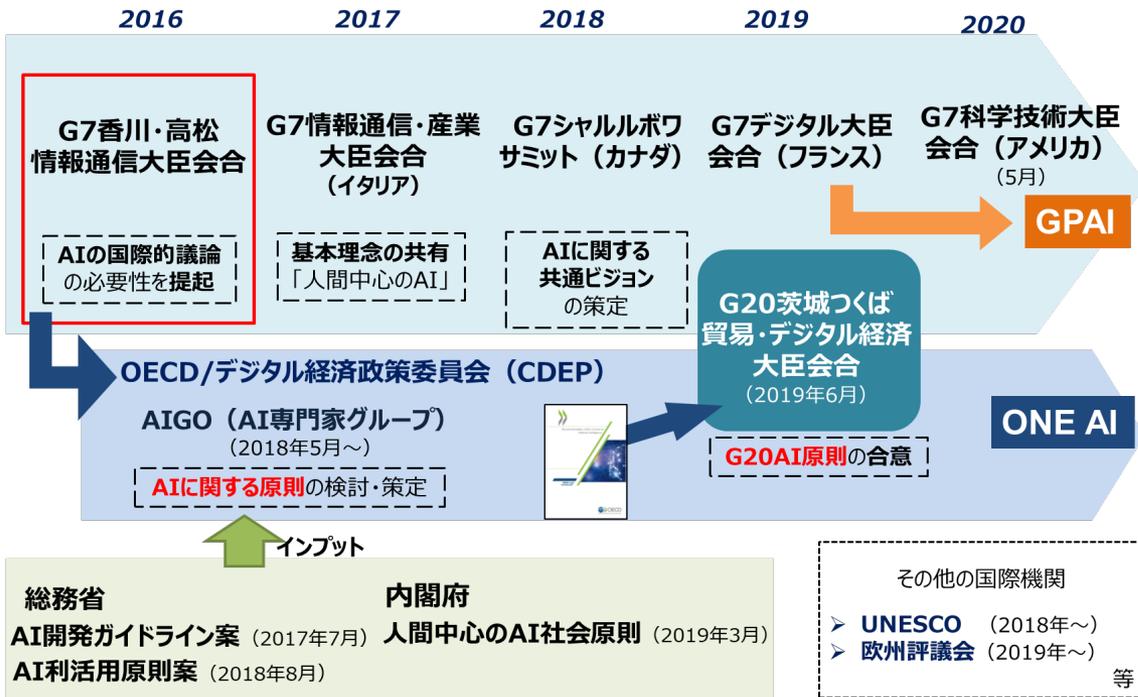
## 機械学習品質マネジメントガイドライン(NEDO、2020)

「内部品質」の向上により「外部品質」を必要なレベルで達成し、最終製品の「利用時品質」を実現すると整理



# 現在進められている議論【総務省資料より】

## A I に関する議論の国内外における大きな流れ



## Global Partnership on AI

- 人間中心の考え方に立ち、「責任あるAI」の開発・利用を実現するために設立された、価値観を共有する政府・国際機関・産業界・有識者等からなる官民間連携組織。2020年6月に創設。
- 2019年ピアリッツサミット (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。
- 参加国 (昨年12月時点) は、18ヶ国+EU: 日本、オーストラリア、ブラジル、カナダ、フランス、ドイツ、インド、イタリア、韓国、メキシコ、オランダ、ニュージーランド、ポーランド、シンガポール、スロヴェニア、スペイン、英国、米国、EU

## OECD Network of Experts on AI

- AIに関するOECD理事会勧告 (2019年5月採択。AI原則及び政府への勧告を含む。) の社会実装の取組の一環として、AIに関する取組の情報共有を進めるためのオンラインプラットフォームである、「AI政策に関するオプザバトリー」 (OECD.AI) に助言を行うAI専門家グループを2020年2月から運用開始。
- メンバーはAI政策の専門家、AI研究者及び技術者、AIに関する法律家及び社会・人文学者など、学際的かつマルチステークホルダーから構成される。
- 須藤教授が参加

## UNESCO

- 2019年11月の第40回UNESCO総会での決定を受け、2020年2月にAI倫理に関する勧告の作成を目的としたアドホック専門家会合を設置。2021年11月 (予定) の第41回総会での勧告の採択を目指す。

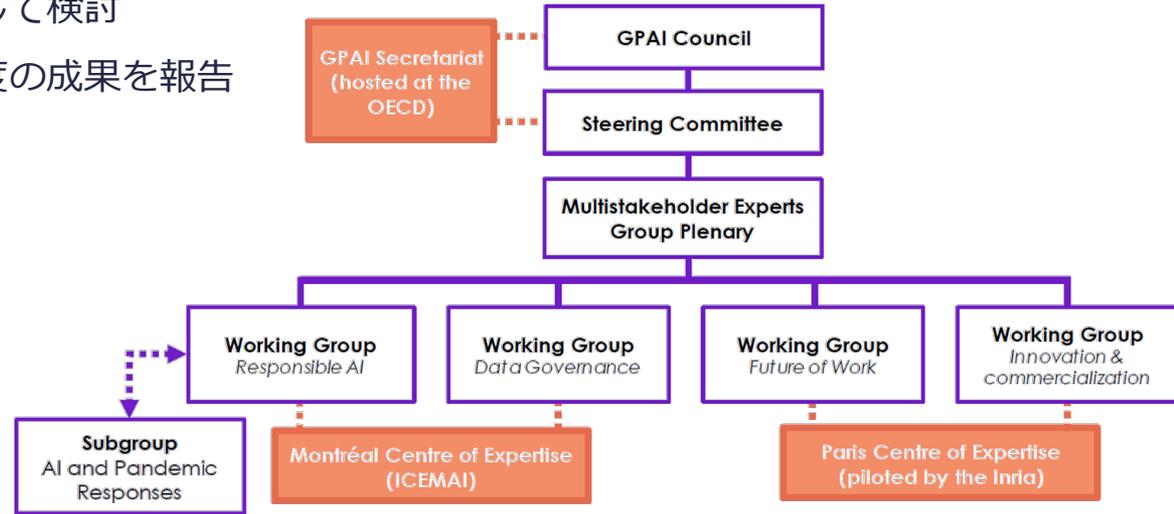
## Council of Europe (欧州評議会)

- 各加盟国において、AIの開発・利用に際する人権・民主主義・法の支配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を主目的とする会合 (AdHoc Committee on AI, CAHAI) を創設。

# GLOBAL PARTNERSHIP ON AI (GPAI)

2020年は4つのWGを設置して検討

12月のPlenary会合で初年度の成果を報告



WG	マンデート
①責任あるAI <b>Responsible AI</b>	SDGsと合致する形で、人間中心のAIシステムの責任ある開発・利活用・ガバナンスを促進する。
②AIとパンデミックへの対応 <b>AI and Pandemic Responses</b>	COVID-19及び将来のパンデミックにおけるAIソリューションの責任ある開発・利活用を促進する。
③データガバナンス <b>Data Governance</b>	人権、包摂性、多様性、イノベーション、経済成長、社会的利益に即すような、AIのデータ収集、利用、共有、蓄積、削除を促進する。
④仕事の未来 <b>Future of Work</b>	AIが労働者に与える影響、仕事の質、包摂性、健康、安全の維持等に関する分析を現在の実例及び将来のビジョンに基づき行う。
⑤イノベーションと商業化 <b>Innovation and Commercialization</b>	民間企業や研究機関がAIイノベーションでの国際連携を推進し、研究成果を製品や製造プロセスに適用、商業化・普及に繋げるための実践的なツールや手法の研究と推奨を行う。

# GPAI MONTREAL SUMMIT 2020



- 2020年12月3～4日にモントリオール専門家センター（ICEMAI）によりバーチャル開催。世界中の主要なAI専門家とGPAI会員の代表者が、GPAIのガバナンス機関であるGPAI評議会と運営委員会の初会合を含むセッション、ネットワーキングの機会、ハイレベルな会合に参加。
- 学界、市民社会、産業界、労働者・労働組合代表、政府、国際機関から200名以上のAI専門家が集まり、GPAIワーキンググループのこれまでの進捗状況をレビューし、今後の作業計画の指針となるフィードバックを提供。



## Video highlights and speeches



Address by Prime Minister of Canada Justin Trudeau  
Introduction by Minister Navdeep Bains, GPAI Council Chair  
(in English and French)



Address by President of France Emmanuel Macron  
(in French with English subtitles)



Watch the GPAI Multi-stakeholders Plenary and address by Plenary Chair Baroness Joanna Shields

# RESPONSIBLE AI WGの初年度の結論

- 全世界から 214のAI原則を収集・分析。参考とすべき30を選択

Table 3: List of 30 Shortlisted Initiatives

Figure 1: Number of initiatives across categories

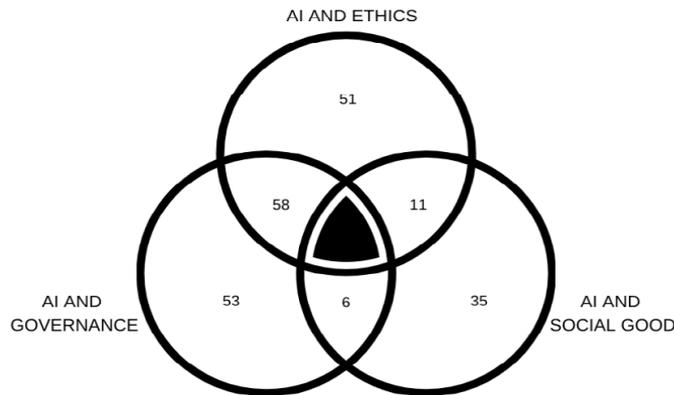
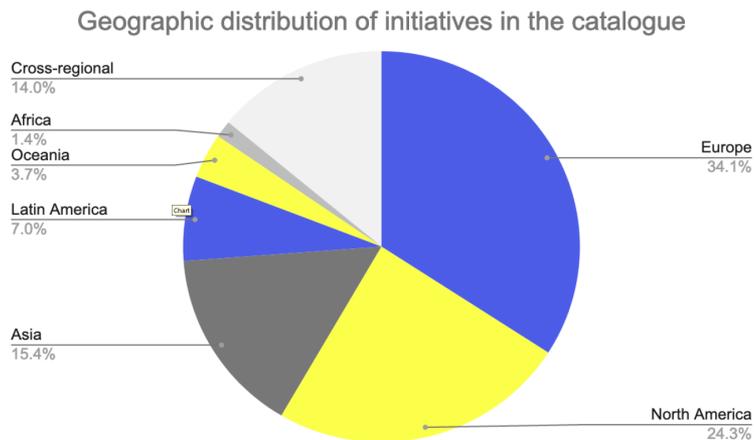


Figure 2: Geographic Distribution of Initiatives in the Catalogue



AI & Ethics	Both AI & Ethics and AI & Governance	AI & Governance	AI & Social Good	Both AI & Ethics and AI & Social Good
Asilomar AI Principles	Algorithm Charter for Aotearoa New Zealand	AI Explainability 360	AI-Based Referral System	HumanE AI Net
CEPEJ Ethical Charter on the Use of AI in Judicial Systems and their Environment	Artificial Intelligence Standardization White Paper	AI Now Report	AI Commons	UNICEF AI for Children
<b>Draft AI R&amp;D Guidelines for International Discussions</b>	Assessment List for Trustworthy AI (ALTAI)	CDEI Review of Online Targeting	AI for Good	
Montréal Declaration: Responsible AI	IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems	Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS)	AI for SDGs Think Tank	
	<b>OECD Recommendation on AI</b>	Global Governance of AI Roundtable	Artificial Intelligence against Modern Slavery (AIMS)	
	Partnership on AI Issue Area on Safety-Critical AI (SCAI)	ISO/IEC JTC 1/SC 42	Artificial Intelligence and Blockchain for Healthcare Initiative in Africa	
	UNESCO Recommendation on the Ethics of AI & AI Decision Makers' Toolkit	<b>Machine Learning Quality Management Guidelines</b>	Elements of AI	
			iGamma	
			Open Kinyarwanda	
			Observatory from the fAIR LAC Initiative	

出典 : The Future Society (December 2020) Areas for Future Action in the Responsible AI Ecosystem.

## RESPONSIBLE AI WGの提言：AREAS FOR FUTURE ACTION

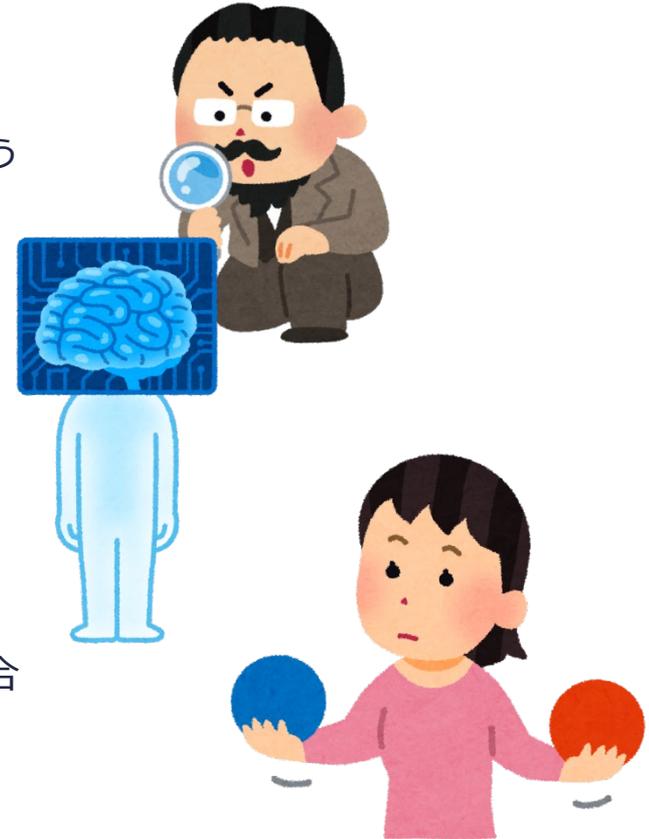
1. 優先課題への資源配分 (Prioritize resources towards the most pressing global issues)
  - 喫緊の課題に確実に取り組むための体系的プロセスを構築
  - 特定された課題に対処するための重点委員会を設置
2. 影響評価を重視したAI原則の設計 (Ensure initiatives are designed for impact)
  - GPAI各国政府間で共通の分類法と国際的な測定システムを開発
  - 広く適用可能で首尾一貫した影響評価方法を構築
3. 変化加速のためにエコシステムを強化 (Strengthen the ecosystem to accelerate change)
  - ガバナンスの課題と手段について重点的に検討する委員会を組成
  - エコシステム内の協調関係を支援
  - 責任あるAIを管理するための政策立案者の能力構築
4. 多様性と包摂性の尊重 (Respect and champion diversity and inclusion)
  - Diversity and Inclusionのグッドプラクティスを開発・普及
  - 代表的な意見を収集するプラットフォームとの戦略的パートナーシップを開始

## RESPONSIBLE AI WG 二年目の現状

- 4つの委員会 (Committee) を組成し、2021年度作業対象を選定中
  1. ガバナンス問題およびその手法 (Issues about and Means of Governance)
    - 各国が直面する多様な環境に応じて多様な解決策があることを念頭におきつつ、AIガバナンス問題についてGPAI参加各国が有する専門知や経験知、さらにはAIガバナンスの現状に関する情報、さらにはGPAIに参加している専門家の知見をプールして相互に有効活用するメカニズムの構築。さらに研究資源に余裕があれば、関連技術の評価、政策の国際比較、キャパシティビルディング、監査手法の研究を実施。【2021/2/26時点の提案】
  2. ソーシャルメディアのガバナンス・透明性 (Governance and Transparency of Social Media)
    - SMS上のコンテンツ推薦に活用しているAIアルゴリズムの透明性を改善の情報収集メカニズムに係る制度的・技術的手法の調査。事業者自身による情報開示ではなく、独立研究者が事業者内部から調査して結果を公表する手法を指向。【2021/2/26時点の提案】
  3. AIと教育 (AI and Education)
    - 教育AIの便益最大化、教育と教師への支援、学習と学習評価の改善、生涯学習の確保、公平・包括的な利用促進、生産・利用者教育、教育への影響の監視など【当初提案】
  4. 創薬とオープンサイエンス (Drug Discovery & Open Science)
    - オープンかつ公平な方法でAIが創薬に貢献できる環境の検討、研究開発の組織化、創薬成果のフェアな分配のためのルール検討【当初提案】

## 【私見】AI導入によって求められる新たな規律

- AIがこれまでの新技術と異なる点は何か？
    - 作動原理がわかりにくい
    - 進歩が速く、かつ、進歩のコントロールがむずかしそう
- ↓
- 差分として求められる規制の追求
    - AIが生み出したものの多くは「改善」
- ↓
- AI導入の限界便益・限界費用の定義・計測が必要
    - 規模×発生確率という尺度
    - 期待・不安と効率性（非効率性）を峻別したうえで統合的に考慮
- +
- 不適切な規制を導入する場合のコストの認識
    - 安全側に振ると過剰規制になり、イノベーションの芽を摘みかねない
    - イノベーションを重視すると過少規制になり、消費者の利益が阻害されるおそれ



## 【私見】 規律デザイン上の配慮

- 実際に規律の運用にあたる民間セクターのインセンティブへの配慮は必須
    - ビジネス上の合理性が見いだせない規律を導入した場合、長期的には維持不可能
    - 研究開発に必要な投資資金を潤沢に集めることができる環境を維持するという視点
  - AIの開発には規模・範囲の経済が強烈に機能する一方、その実施はボーダレスに実行可能であるという特性への配慮
    - 大量の学習用データを早期に集積した主体が他を圧倒できる
    - データやソフトウェアは容易に国境を超える
      - データ等の移転制限はAIの生産性に悪影響
    - AI開発がもたらすマーシャルの外部効果を考えると、国家間の協調が必須
      - Regulatory shopping、Race to the bottomの阻止
- 
- 民間ビジネスとの整合性を確保した規律を国際協調の下で導入する努力
    - Responsible AIのメリットを一般利用者に周知する努力
    - AI dividendの分配を国境を越えてフェアなものとする努力

## 【私見】 マルチステークホルダー、共同規制、情報開示、ベータ版

- 政策担当者が直面する情報の非対称性を抑制するとともに、市場参加者にとっての不確実性を抑制することが何よりも重要
  - 情報の非対称性を抑制するためには、マルチステークホルダーの関与、さらには共同規制の採用が効果的
  - 不確実性抑制のためには政府において進められている関連議論についての情報開示が有用
  
- そのうえで、AIは発展途上の技術であることに対する認知を高める努力をリテラシーの高い側（供給側に限らない）が継続する必要がある
  - ベータ版であることへの理解
  - 「機械学習」の理解
  - その他、最新研究成果への理解

