AIの倫理、日本発AI諸原則／ガイドライン

中央大学　国際情報学部 教授&学部長
米国弁護士（NY州）　平野（ひらの）　晋（すすむ）

1. はじめに

日本や各国の人工知能（AI）推進政策においては、〈STEM（ステム）：Science, Technology, Engineering, and Mathematics〉の頭字語に代表される理数工学系教育の重要性が専ら指摘されている。しかし社会に広まっているAIの不安が払拭されなければ、AIは普及しない。今後目指すべきAIは、「人間中心の」とか「信頼できる」等の修飾語が「AI」の前に付記される〈条件付きのAI〉であるとされる。そこで必要になるのが、AIの倫理的・法的・社会的影響（ELSI（エルシー）：Ethical, Legal, and Social Implications）を検討した上で必要となる規範や施策である。AIの倫理規範や諸原則やガイドライン等が世界中の公的機関やAI関連組織において検討され、かつ公表されている事実に鑑みれば、ELSIの重要性はもはや説明するまでもないであろう。

ところで本稿は、2020年6月23日に開催された第88回情報通信研究会（リモート開催）に基づいている。本稿でも、日本内外の有識者会議（下記）においてAIの規範構築に携わってきた筆者の経験に基づいて、国内外のAI諸原則（Principles）やガイドライン（指針）の構築発展史について、記しておこう。

- 経済協力開発機構（OECD）「AI専門家会合」（AIGO[*1]：Artificial Intelligence expert Group at the OECD）2018年9月～2019年2月
- 内閣府「人間中心のAI社会原則［検討］会議」2018年4月～2019年3月
- 総務省「AIネットワーク社会推進会議」2016年10月～現在
- 総務省「AIネットワーク化検討会議」2016年1月～6月

2. AI諸原則／ガイドライン発展史：いかに日本のAI規範が国際標準に影響を与えたのか

2019年5月には、日本及び主要国が参加する国際機関のOECDが、いわゆる〈OECD_AI原則〉についての理事会勧告を発表した[*2]。翌6月には、日本が議長国を務めたG20貿易・デジタル経済大臣会合が、そのOECD_AI原則を、いわばコピペする形で「G20_AI原則」を採択している[*3]。そして、OECD_AI原則の形成には、日本が国際社会に向けて採用を働き掛けた総務省や内閣府の有識者会議がまとめたAI諸原則／ガイドラインが大きく影響を与えたと評されている[*4]。その影響過程を以下で紹介してみよう。

2.1　総務省「AIネットワーク化検討会議」（2016年1～6月）と、香川県高松市開催のG7情報通信大臣会合（2016年4月）におけるAI開発ガイドライン検討の提案

筆者が現在所属する中央大学・国際情報学部（iTL[*5]）の須藤修教授（当時は東大大学院・情報学環教授、東大・総合教育センター長）が座長を務め、やはりiTL所属の石井夏生利教授（当時は筑波大学・図書館情報メディア系准教授）が構成員で、かつ筆者（当時は中大・総合政策学部教授、同大学院総合政策研究科委員長）が座長代理を務める総務省「AIネットワーク化検討会議」が、2016年1月から「AI開発原則」のたたき台的な案（8原則）を検討していたところ、同年4月にG7（先進7か国・地域）情報通信大臣会合が高松市で開催され、ホスト国である高市早苗総務大臣（当時）から、そのたたき台を紹介しつつOECDの協力も得て国際的なガイドラインの策定を提唱したところ、参

---

*1　AIGOの会合に出席した筆者の経験に基づけば、〈AIGO〉は「エイゴー」または「エイ・アイ・ゴー」と発音する。

*2　OECD「42か国がOECDの人工知能に関する新原則を採択」http://www.oecd.org/tokyo/newsroom/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence-japanese-version.htm（last visited Sept. 13, 2020）.

*3　G20 Ministerial Statement on Trade and Digital Economy, 3-4, 11-12 https://www.soumu.go.jp/main_content/000625715.pdf（last visited Sept. 13, 2020）.

*4　総務省・AIネットワーク社会推進会議「報告書2019」2頁 https://www.soumu.go.jp/main_content/000637096.pdf（last visited Sept. 13, 2020）.

*5　中大・国際情報学部は、教育・研究の柱を〈IT〉と〈Law：法学〉の2つに置いているので〈IT+Law=iTL（アイ・ティー・エル）〉と呼ばれる。〈i〉を、あえて小文字にして強調することで、〈integration〉等の意味も含意させることにより、ITと法学を統合させた学際法学を目指している。

■図1. 総務省・AIネットワーク化検討会議から香川県高松市G7情報通信大臣会合へ

加国大臣達から賛同が得られた[*6]（図1参照）。

なお、「AIネットワーク化検討会議」は発展的に改組されて、2016年10月から新たに「AIネットワーク社会推進会議」が活動を開始し（現在も活動継続中）、その議長を引き続き須藤修先生が務められ、筆者は幹事及び開発原則分科会長を務めることになり、前記AI開発原則たたき台を更に洗練化させ、かつこれに解説も付けたガイドラインを「国際的な議論」に供するために、起案する役目を引き受けた[*7]。

2.2 国際社会への働き掛け

AIネットワーク社会推進会議が活動を開始した翌月の2016年11月17日には、総務省の同会議におけるAI開発原則／ガイドラインの検討状況を広報すべく、筆者はパリに出張して、「OECD技術予測フォーラム」において発表を行った[*8]。2か月後の2017年1月12日にはワシントンD.C.に筆者が出張し、在米日本大使館が後援してカーネギー国際平和基金において開催されたカンファレンスにおける基調講演として、日本の取組みを紹介[*9]。その2か月後の2017年3月13日と14日には、総務省が主催する国際シンポが東京大学伊藤謝恩ホールにおいて開催され、筆者もAI開発ガイドライン案等の検討状況を発表した[*10]。その4か月後の2017年10月26日と27日には、OECDと総務省が共催したカンファレンスに参加すべく、再びパリに出張して、AI開発ガイドライン等の検討状況を発表し、かつパネルディスカッションにも

＊6 「報告書2019」*supra* note (4), at 1-2頁. *See also* 総務省「経済協力開発機構（OECD）における人工知能（AI）に関する取組」https://www.soumu.go.jp/menu_seisaku/ictseisaku/oecd_ai/index.html (last visited Sept. 15, 2020).

＊7 「資料1 AIネットワーク社会推進会議 開催要綱（案）」https://www.soumu.go.jp/main_content/000447310.pdf (last visited Sept. 14, 2020).

＊8 OECD, *Technology Foresight Forum 2016 on Artificial Intelligence (AI)*, http://www.oecd.org/sti/ieconomy/technology-foresight-forum-2016.htm (last visited Sept. 14, 2020).

＊9 Carnegie Endowment for International Peace, *Artificial Intelligence and U.S.-Japan Alliance Engagement*, Jan. 12, 2017, https://carnegieendowment.org/2017/01/12/artificial-intelligence-and-u.s.-japan-alliance-engagement-event-5464 (last visited Sept. 14, 2020).

＊10 総務省・国際シンポジウム「AIネットワーク社会推進フォーラム」https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02iicp01_04000093.html (last visited Sept. 14, 2020).

■図2. 総務省による国際社会への働き掛け

参加[*11]。加えて、OECDのDouglas Frantz事務次長（当時）等とのインタビュー動画にも出演した[*12]（図2参照）。

以上のような、欧米に対する文字通り矢継ぎ早で、かつ積極的な日本の取組みの広報活動が、その後OECD_AI原則が日本の主張を反映することにつながったものと思料される[*13]。

2.3 内閣府「人間中心のAI社会原則［検討］会議」と、AIGO（OECD_AI専門家会合）

OECDと総務省が共催したカンファレンスから約半年後の2018年5月には、内閣府内に「人間中心のAI社会原則検討会議」（後に「人間中心のAI社会原則会議」に改名[*14]）が活動を開始し、須藤先生が議長を務め筆者が構成員として参加して、AI社会原則を構築の上公表した[*15]。そして同会議の開催中に、OECDはAIの理事会勧告に向けて、世界の専門家を集めた「AI専門家会合」（AIGO：Artificial Intelligence expert Group at the OECD）を開催することになり[*16]、須藤先生と筆者が日本代表としてそこに参加することとなった[*17]。その初会合はパリのOECDにおいて、2018年9月に開催され筆者が参加し、僅か2か月後の11月に再びパリのOECDで開催された2回目の会合には須藤先生と筆者が参加。そこから再び僅か2か月後の2019月1月にボ

*11 総務省「報道資料 AIに関する国際カンファレンスの開催（OECD・総務省共催）」平成29年10月20日 https://www.soumu.go.jp/menu_news/s-news/01tsushin08_02000090.html (last visited Sept. 14, 2020).

*12 OECD, *Intelligent Machines, Smart Policies：Interview*, https://www.facebook.com/watch/?v=10154858561222461 (last visited Sept. 14, 2020).

*13 「報告書2019」*supra* note (4), at 2頁.

*14 内閣府・統合イノベーション戦略推進会議決定「人間中心のAI社会原則」13-14頁（平成31年3月29日）https://www8.cao.go.jp/cstp/aigensoku.pdf (last visited Sept. 14, 2020).

*15 *Id. See also* 内閣府「人間中心のAI社会原則検討会議」https://www8.cao.go.jp/cstp/tyousakai/humanai/index.html (last visited Sept. 14, 2020).

*16 *See* OECD, *OECD Creates Expert Group to Foster Trust in Artificial Intelligence*, Sept. 13, 2018, https://www.oecd.org/going-digital/ai/oecd-creates-expert-group-to-foster-trust-in-artificial-intelligence.htm (last visited Sept. 14, 2020).

*17 OECD, List of Participants in the OECD Expert Group on AI (AIGO), https://www.oecd.org/going-digital/ai/oecd-aigo-membership-list.pdf (last visited Sept. 14, 2020).

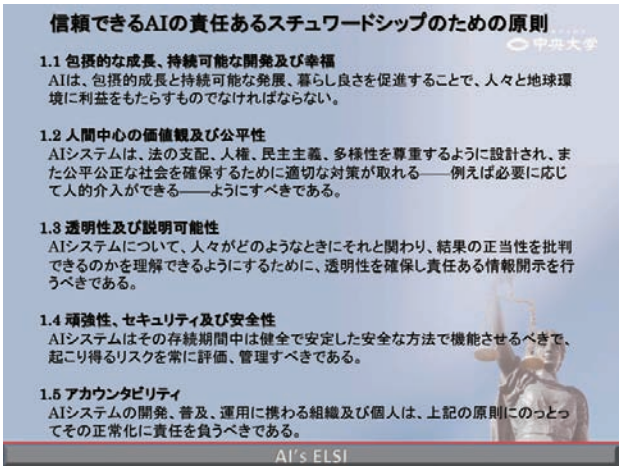
■図3. 内閣府「人間中心のAI社会原則［検討］会議」からAIGO（OECD_AI専門家会合）へ

ストンのマサチューセッツ工科大学で開催された3回目の会合にも須藤先生と筆者が参加し、これらの会合では総務省の取組みに加えて、前記内閣府のAI社会原則の検討状況もタイムリーに紹介して、OECD_AI原則に反映するように働き掛けた（**図3**参照）。

3. OECD_AI原則

2019年5月には、OECD_AI原則が理事会勧告として採択された[*18]。その概要は**図4**のとおりであり、日本のAI諸原則やガイドラインと整合性のあるものとなった。詳細は拙書『ロボット法:AIとヒトの共生に向けて』281～313頁（2019年）を参照いただきたい。

なお最近、OECDは「OECD.AI：Policy Observatory」[*20]というウエブサイトを公開していて、本稿執筆時点では、その中の「AI Principles」というアイコン部分をクリックして現れる解説部が[*21]、AI専門家会合（AIGO）における議論をよく反映しているので、読者の参考になろう。



■図4.「OECD_AI原則」の概要[*19]

OECD_AI原則が、その僅か1か月後には、日本で開催されたG20の貿易・デジタル経済大臣会合において、いわゆるコピペするように採用されて「G20_AI原則」と称され

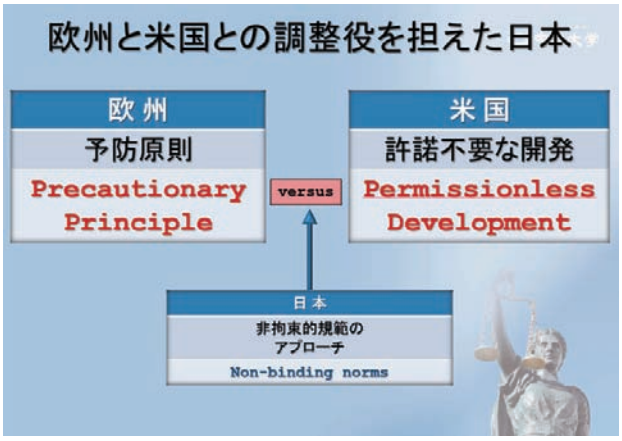
＊18　OECD「42カ国」*supra* note (2).
＊19　*Id.* 及び「人工知能に関する理事会勧告」（総務省和訳版）8～10頁 https://www.soumu.go.jp/main_content/000642217.pdf (last visited Sept. 14, 2020) から引用.
＊20　https://www.oecd.ai/ (last visited Sept. 14, 2020).
＊21　https://www.oecd.ai/ai-principles (last visited Sept. 14, 2020).

たことは、前述したとおりである[*22]。

4. 勝因

ところで、日本発のAI諸原則／ガイドラインの考え方が、OECD_AI原則に採用された勝因は何であろうか。私見では、主な勝因は以下であると思料しているので、披露しておこう。

- 〈欧州〉対〈米国〉の対立構造の中で、日本が中庸な代替案を提言できたこと
  - 日本の非拘束的アプローチが受容されたこと（**図5**参照）。



■図5. 欧米の架け橋を担った日本

- マルチ・ステイク・ホールダーなアプローチが受容されたこと
- 国際社会に対して、時宜を逸することなく、間髪を入れず矢継ぎ早かつ積極的な広報活動を続けたこと

5. おわりに

AIは、各国や経済界が経済発展の起爆剤として期待の大きい科学技術である。しかし他方では、人々の職場や生活に負の影響も与えるという不安が広がっている[*23]。その不安は根拠がないと主張する者も、残念ながら一部見受けられるけれども[*24]、筆者が多くの有識者会議と国際会議を通じて理解したところによれば、実際にAIには欠点や、治癒・改善すべき点や、取り組むべき問題点も多く見受けられた[*25]。したがって、AIの開発者や利用企業等の関係者は、人々が抱く不安を真摯に受け止め、かつAIの欠点をいたずらに正当化するのではなく、寧ろ改善に努めるべきであろう。特にOECD_AI原則が「信頼できるAI」（trustworthy AI）という文言を用いて、信頼の重要性を強調している点は、関係者による改善の努力を求めている意思の表れであるから、その意思を誠実に受け止めた上で、日本の関係者の間でもAI諸原則・ガイドラインが実装されて、より良いAIの開発、利活用、等々が目指されることを期待したい。

［2020年6月23日　情報通信研究会より］

＊22　*See supra* text accompanying note (3).

＊23　*E.g.,* 内閣府・統合イノベーション戦略推進会議決定「AI戦略～人・産業・地域・政府全てにAI～」53頁（令和元年6月1日）https://www.kantei.go.jp/jp/singi/ai_senryaku/pdf/aistratagy2019.pdf (last visited Sept. 14, 2020).

＊24　特に汎用AIをソフトローの射程に入れる姿勢に対して、残念ながらいまだ一部に強い抵抗が見受けられる。ELSI教育の重要性を強く感じざるを得ない現象である。

＊25　例えば、制御不可能性、不透明性（ブラック･ボックス問題）、差別的な判断等々の欠点や、汎用AIへの懸念が指摘されている。汎用AIもソフトローの射程に入れるべきという有識者会議の見解とその理由については、see, *e.g.,* 総務省・AIネットワーク社会推進会議「報告書2017」26～28頁＆脚注63～67（平成29年7月28日）https://www.soumu.go.jp/main_content/000499624.pdf (last visited Sept. 15, 2020)；「AIネットワーク社会推進会議 第6回議事概要」（平成29年7月25日）https://www.soumu.go.jp/main_content/000513654.pdf (last visited Sept. 15, 2020).