

## 第7回 人間中心のAI社会原則検討会議議事録

1. 日 時 平成30年11月6日(火) 10:00~12:00

2. 場 所 中央合同庁舎第4号館 第1特別会議室

### 3. 出席者

議長	須藤	修	東京大学大学院情報学環教授 東京大学総合教育研究センター長
副議長	北野	宏明	一般社団法人日本経済団体連合会未来産業・技術委員会AI活用原則TF主査 株式会社ソニーコンピュータサイエンス研究所 代表取締役社長
構成員	安宅	和人	ヤフー株式会社CSO
	同	岩本	敏男 株式会社エヌ・ティ・ティ・データ相談役
	同	浦川	伸一 損害保険ジャパン日本興亜株式会社取締役常務執行役員
	同	江間	有沙 東京大学政策ビジョン研究センター特任講師
	同	大屋	雄裕 慶應義塾大学 法学部教授
	同	木俣	豊 情報通信研究機構知能科学融合研究開発推進センター長
	同	近藤	則子 老テク研究会 事務局長
	同	関口	智嗣 産業技術総合研究所理事
	同	高原	勇 トヨタ自動車株式会社BR-未来社会工学室長 筑波大学未来社会 工学開発研究センター長、特命教授
	同	武田	晴夫 株式会社日立製作所理事 研究開発グループ技師長
	同	中川	裕志 理化学研究所革新知能統合研究センターグループディレクター
	同	新居	日南恵 株式会社manma代表取締役社長
	同	羽鳥	裕 公益社団法人日本医師会常任理事
	同	樋口	知之 統計数理研究所長
	同	平野	晋 中央大学総合政策学部教授 大学院総合政策研究科委員長
	同	福岡	真之介 西村あさひ法律事務所弁護士
	同	堀	浩一 東京大学大学院工学系研究科教授
	同	丸山	宏 株式会社Preferred Networks PFNフェロー

関係府省	赤石	浩一	内閣府政策統括官（科学技術・イノベーション担当）
	同	佐藤	文一 内閣府政策統括官（科学技術・イノベーション担当）付審議官
	同	新田	隆夫 内閣府政策統括官（科学技術・イノベーション担当）付参事官
	同	今川	拓郎 総務省情報流通行政局情報通信政策課長
	同	井上	知義 総務省情報通信政策研究所長
	同	原	克彦 文部科学省研究振興局参事官（情報担当）
	同	江浪	武志 厚生労働省大臣官房厚生科学課医療イノベーション企画官
	同	松本	真太郎 経済産業省産業技術環境局研究開発課産業技術プロジェクト推進室長
	同	檜山	洋平 国土交通省大臣官房技術調査課課長補佐
	同	菊池	健秀 国土交通省総合政策局技術政策課専門官
ワグサーハ	八山	幸司	内閣官房情報通信技術（IT）総合戦略室参事官
	同	庄崎	未果 内閣官房副長官補日本経済再生総合事務局企画官
	同	吉田	恭子 内閣官房内閣サイバーセキュリティセンター内閣参事官
	同	檜尾	洵 内閣府知的財産戦略推進事務局参事官補佐
	同	高倉	秀和 内閣府宇宙開発戦略推進事務局参事官
	同	佐々木	亨 総合海洋政策推進事務局参事官
	同	三原	祥二 個人情報保護委員会事務局参事官

#### 4. 議題

- （1） 国内外の動向
- （2） とりまとめ構成案について

#### 5. 配布資料

- 【資料1】 国内外の動向
- 【資料2】 今後のスケジュール

- 机上資料1 人間中心のA I 原則検討会議とりまとめ一次案  
机上資料2 北野副議長提出資料  
机上資料3 第6回人間中心のA I 社会原則検討会議議事録(案)

## 6. 議事

### 【須藤議長】

定刻になりましたので、第7回人間中心のA I 社会原則検討会議を開催いたします。

本日はご多忙の中お集まりいただきまして、ありがとうございます。

今朝の朝刊にありましたように、EUがかなり規制色の濃いA I 倫理指針を提案しているようです。まだ正式発表ではないということですが、日経が1面トップで書いておりますが、恐らく、来週月曜日に開催されるOECDの会議には、EUはこの案のもう少し詳細なものを正式にEUの意見として出してくると思います。

本会議としては、これについては今ではなく、もう少し国際的な動きがあった後に、正式に議論はしたいと思います。これは今後のこの報告書案にも影響を及ぼす可能性があると思います。というのは、EUの動きはG A F Aあるいは中国との緊張関係で議論されることになると思いますので、そこら辺は考慮しますが、今見た感じ、今朝ニュースを見た感じでも少々、かなり強い言い回しがありましたので、慎重に議論しなければならないと思います。

研究開発あるいは活用を阻害しないようにも考えないといけないと、私は個人的には考えています。と同時に、きちんと社会的なガバナンスがきくようにする、その両立が求められるだろうと思います。

それでは、これより議事に入りますが、その前に配付資料の確認を事務局よりお願いいたします。

### 【新田参事官】

事務局を務めます内閣府の新田でございます。よろしくお願いいたします。

本日の配付資料でございますが、資料1国内外の動向、資料2今後のスケジュール、机上配付の資料として、机上資料1が前回から再構成いたしました取りまとめ構成案、机上資料2として、北野副議長から提出いただいておりますパワーポイントの資料、それから、前回議事録案を机上資料3ということで配付してございます。

過不足等ございましたら、事務局までお願いいたします。

**【須藤議長】**

どうもありがとうございます。

本日の会議は公開で進めたいと思います。

早速ですが、議題1の国内外の動向ということで、9月のOECDにおけるAIに関する専門家会合とEUの動向について、事務局からまとめてお話を頂いて、その後、OECDの専門家会合にご出席されました平野構成員から、簡単にOECDの議論をご報告いただきたいと思います。

まずは事務局から、お願いいたします。

**【新田参事官】**

それでは、お手元資料の1をご覧ください。

国際的な議論の動向ということで、2点お示しさせていただいてございます。

1点目はOECDの関係で、去る9月24日、25日にAIに関する専門家会合、AIGOと申しますが、その第1回会合が開催されております。

こちらは2019年中の理事会勧告策定を視野に入れて、その検討をするものでございます。この理事会勧告の中では、AIの信頼構築と社会実装を促す原則が策定され、OECDとして盛り込む予定となっております。

先ほど、議長からもございましたが、日本からは須藤議長と平野構成員がご参加でございます。

今後、第2回会合は来週11月12日、パリ、第3回会合は翌年1月16日から17日、ボストンという予定となっております。理事会勧告の策定プロセスが着実に進んでいるという状況でございます。

2点目は欧州連合のAI倫理に関する議論の動向でございます。先ほど、日経新聞の記事のご紹介を須藤議長からも頂きましたが、欧州連合におきましては、今年4月にAI倫理を策定するための専門部会を立ち上げ、AI for Europeというドキュメントが出ておまして、そちらの資料を抜粋したものでございます。

プロセスとしては、今年12月にドラフトを策定、最終版を来年3月に策定するというところで、このAI倫理原則検討会議のタイムフレームと非常に似通っているという状況でございます。

ちなみに、ここに来年の中頃にプロダクト・ライアビリティ・ディレクティブということで、EUにおきます製造物責任指令に関する解釈に関するガイダンスの文書を策定する予定となっております。

そのほか、説明可能なAI開発支援などもフレームワークとして盛り込まれているということと、下に\*で示しておりますように、ガイドラインのドラフト案の中には、この4月のドキュメントの段階では、公平性、安全性、セキュリティ、社会的包摂性、アルゴリズムの透明性、更に広範囲な議論として、プライバシー、尊厳、差別禁止という基本権利への影響というところも盛り込まれる予定であります。先ほど、AI for peopleという記事のご紹介がございましたが、この議論とおそらく、AI for peopleという有識者会議の議論を土台に、今後、EUにおけるAI倫理の議論、策定が進んでいくと思われま。

本会議と同じようなタイムフレームで今後進んでいきますので、国際的な情報発信を我々の方も想定しておりますところ、諸外国の議論の動向も十分見据えながら、柔軟に議論をしていければいいのかなと思います。

本日、取りまとめ一次案として議論のベースをお示ししておりますが、諸外国の動向などもしっかり確認しながら議論を進めていければと考えているところでございます。

事務局からは以上です。

#### 【須藤議長】

どうもありがとうございます。

それでは続きまして、平野構成員、ご報告をお願いいたします。

#### 【平野構成員】

中央大学の平野でございます。

今ご報告いただきました資料1の1ページ、最初のブレットポイントのパラグラフの2行目でございます。

OECDにおいて理事会勧告に向けて今やっている、2段落目のこちらは、「エイゴ」と発音するようで、Artificial Intelligence expert Group at the OECDの略でAIGOということです。

最初のパラグラフの2行目に戻りますが、理事会勧告に向けて文言をつくっていかなくてはいけないということで、AIGOはその作業部会というか、文言を詰めていく、そういう会合

でございます。

目的は1段落目の2行目最後、To Foster Trust in A Iとadoption of A Iに向けて原則をつくっていかうじゃないかということでございます。

東京大学の須藤議長と私がそこに入っております、9月の議論に私が参加してきたところでございます。日本での動きも、彼ら相当勉強しております、休憩時間に議長のウォンキー・ミン氏、その下で働いているアン・カブラン課長から、非常に日本は貢献していただいておりますというふうなお話も賜っております。

今後のスケジュールが3番目の段落でございまして、来週の月曜日にパリで朝から夕方過ぎまで終日この文言の議論、そして予定では1月に、今度はボストンのM I Tの場所を借りて、そこでA I G Oの会合を続けてやるとのスケジュールになってございます。

#### 【須藤議長】

どうもありがとうございます。

A I G OのメンバーにはM I Tの教授も参加されております。

M I Tは今度新しいカレッジを立ち上げて、1,000億円ぐらいのファンドを集めたと、I B Mや中国の企業からもファンドが出ているということは聞いています。

と同時に、これは私が仲介の仕事をしたのですが、東京大学とM I TがA I、ロボティクスの関係を今後強化しようということで、日経とソニーの共催で、3月に1週間にわたってブートキャンプを開催するというのを決定しています。これは、世界各国から選ばれた若手経営者とM I T、若手研究者が徹底的に討論をすると。共同のプロジェクト等が打てるのであれば打っていく。と同時に、M I Tはファンドを固めるということになります。

グーグルの元CEOのシュミット（現、グーグルの持株会社米アルファベットの取締役兼顧問）は、現在、M I Tに所属して、世界中を歩いて、ファンドを形成するためにご活躍になっています。

東京大学としても、M I Tとの連携を強化して、今年は協力ですが、来年は共催に持って行って、我々もファンドをM I Tとともに固めていく作業に移ろうということで、今、準備に入っているということです。

こういう大学、産官学の連携は、今回の取りまとめ案にも出てきますが、かなりいろいろなバリエーションがあつていいと思いますし、活発にそれが行えるように、規制などで邪魔にな

るようなものは取り払っていかないといけないし、一方で、利用者の立場というのをきちんと守っていくという、ガバナンスも同時に満たさなければならないだろうと思います。

どうもありがとうございます。

それでは、ただいまのお二人のご報告について、何かご質問等があればご発言ください。いかがでしょうか。

#### 【武田構成員】

質問というよりも補足ですが、今後のスケジュールの資料にユネスコのハイレベル会合が1月にあるとありますが、その前の12月に全アフリカのユネスコ主催のA Iのフォーラムがございます。国際的に結構大事な場だということで、日本も何とかそこで発信すべきではないかという話がございます、僭越ですが、私が、参加してまいります。今日の会の議論を踏まえて、何とかアフリカ諸国が日本を向いて、日本の意見に同調してくれるような下地をつくってまいります。ご支援よろしく申し上げます。

#### 【須藤議長】

どうもありがとうございます。

私も武田構成員にご出席いただけるということを政府の方から聞いております。かなり重要な役割になると思います。よろしくお願い申し上げます。

ほか、何かございますでしょうか。

#### 【安宅構成員】

安宅です。

先日、EUでブリュッセル事務所長されている井上博雄さんとお会いしたところ、EUの日本に対するまなざしは、かつてないほど熱く親和的であるとお聞きしています。歴史始まって以来じゃないかというぐらいということなので、この機を逃すことなく、しっかり手を組めるところは大きいかなと思っています。是非よろしく申し上げます。

#### 【須藤議長】

どうもありがとうございます。

北野副議長、どうぞ。

**【北野副議長】**

民間レベルですが、Partnership on AIの総会、メンバーシップミーティングが来週、サンフランシスコでございます。新たに2つのワーキンググループが立ち上がります。一つは”Social and Societal Influences of AI”、もう一つは”Collaborations Between People and AI Systems”で、”Collaborations Between People and AI Systems”の方は私がco-chairを務めることになりました。正に人間中心AIの一番中核のワーキンググループになると思いますので、そこではいろいろなことを発信していきたいと思います。

また、昨日、経団連の方からSociety 5.0の提言書が出まして、ウェブでもダウンロードできますので、ご覧いただければと思います。ダイバーシティとサステナビリティを軸とした価値観で、デジタルトランスフォーメーション、AI-Ready化を進めようという内容になっております。

**【須藤議長】**

いろいろな動きがございます。ありがとうございます。

安宅構成員から今お話がありましたEUの動きですが、やはり個人情報保護法制で、オムニバス法制で体系的なものがあるということで、GDPRとの連携が可能とEUは判断、これを日本側（がわ）で頑張って対応していただいたのは、堀部先生を委員長とする個人情報保護委員会でした。非常にいい関係を築かれました。

一方、米国はEUとのそのような関係は築けていないということで、日本は今、EUと日本のデータを相互移転は可能という枠組みで議論していますが、米国は今まだできないという状況にあるということになります。

そのように、今後グローバルな連携もかなりあり得ると思いますので、ここら辺も皆様の情報あるいはご活動をお願い申し上げたいと思います。

**【赤石統括官】**

議長、すみません、事務局から1点。

**【須藤議長】**

はい、どうぞ。

#### 【赤石統括官】

EUとの関係の補足ですが、10月22日にEUの副大統領級のカタイネン副委員長が来日され、外務大臣と経産大臣との間で議論しています。貿易問題からデジタル問題まで含めて議論しているのですが、ハイレベルでもAIの原則については協力していこうという話になっておりますので、先ほど、事務局から申しました通り、全くEUの動きを無視してということではなくて、お互いによく歩調を合わせながら、一体となってメッセージを出していくと。対中国との関係でも、アフリカとの関係でも、OECDとの関係でも、国連との関係でも、こういった動きをとっていくことが重要だと政府は考えておりますので、是非協力をお願いできればと思います。

#### 【須藤議長】

赤石統括官、ありがとうございます。

今おっしゃっていただいたように、本会もEUとの歩調も視野に入れて構成案を取りまとめていくことが重要になろうと思います。皆様、何とぞよろしくお願いいたします。

それでは、議題2に入りたいと思います。

まず、当初の予定では、前回の取りまとめ構成案に基づきまして、少人数のチームによる執筆作業に入る予定でしたが、取りまとめ構成案の再構築を事務局が行ってまいりました。まずは再構築されたAI原則会議取りまとめ構成案を事務局からご説明いただきたいと思います。

それでは、事務局、よろしくお願いいたします。

#### 【新田参事官】

机上資料1「人間中心のAI原則検討会議とりまとめ 一次案」をご覧ください。

2ページ目、「はじめに」でございます。要素を少し追加しておりますが、大きな流れは変わっておりません。

これまでの歴史を振り返ると、技術のブレークスルーで大きな社会の変革があった。今、国際的には地球規模の課題がある。国内においてもさまざまな社会課題がある中、AIが技術のブレークスルーをもたらすものとして、こういった地球規模あるいは国内の社会課題の解決に

貢献し得るということ。そのためにもSociety 5.0のモデル国家に日本がいち早くなっていく必要があると。そのためにはA I -Ready化が必須であり、あらゆる社会システムをA I -Readyにしていかなければならないということをございまして、この文書は「A I -Readyな社会」を実現する上で必要なA I 社会原則について議論するものである、という前書きとなっています。

2 ページ目の最後の段に脚注といたしまして、本会合においては、「A I」という言葉を、特定の技術を指すものではなくて、広く先進的な情報システム一般を指すものとして捉え、議論するという、これまでのこの会合のディスカッションを踏まえて脚注を追加することをご提案差し上げたいので、この点についてはまたご議論いただければと思います。

3 ページ目にはSociety 5.0の基本理念ということで、前回の構成案では、基本理念はハピネスというアウトカムを目指しながら、その要素としては5つ提示させていただいたと思います。その後、議論の中で、国連でもよく議論になるディグニティ、人間の尊厳はあってもよいのではないかなどのご意見もあり、少し集約を試みしましたので、この辺もご確認いただければと思います。

1 番目のブレットですが、Society 5.0の最上位のアウトカムは、A I とデータが多様な人々の幸せ、ハピネスをサポートし、よりよい社会を構築するということであるとしております。

このA I による社会の変革というのは、社会のありようを大きく変化させるものでありますが、今後出現する社会というのは以下の3つの価値が尊重されるもの、ということで、1つ目に議論もありましたディグニティ、2つ目に多様な背景を持つ人々が多様な幸せを追求できる社会ということでダイバーシティ・アンド・インクルージョン、3つ目に持続性ある社会、サステナビリティということで、頭の整理としては、ディグニティは比較的個人、社会を構成する個々人の尊厳を尊重すると、それは、人間の集まりである社会レベルになった多様な人々が多様な幸せを追求できるという②、それから、時間軸を加えて持続的である必要があるということで持続性ある社会と、頭の整理としてはそんな感じになるのかなと考えております。

ディグニティは、前回深く議論していなかったと思いますので、改めて簡単にご説明したいと思いますが、簡単に言いますと、単にA I で効率性や利便性が実現できるということではなく、また逆に人間がA I に依存したりするようなことになってはよくないと。人間がA I を便利な道具として使いこなすことで、多様な能力を発揮できるとか、創造性を発揮できるとか、やりがいのある仕事に従事できるとか、そういうふうな社会である必要があるというのがディ

グニティ。

2 番目、ダイバーシティ・アンド・インクルージョン。これは前回とあまり内容としては変えておりませんが、こういったインクルージョンというのが現代の国家における一つの理想であり、大きなチャレンジであるといったことを書いています。

3 つ目、持続性ある社会ということで、これは、いわゆる地球規模の課題でございます気候変動、格差、生物学的多様性の喪失、資源の逼迫（ひっばく）といった、地球規模のサステナビリティの脅威になっているものに対する科学的・技術的蓄積を A I で強化して、持続性ある社会をつくる必要があるということと、また、特に格差に関係して、一極集中による地域格差、A I スキルの差による所得格差といったことは、社会全体の崩壊につながってしまうおそれがありますので、A I の活用が社会の格差解消にむしろ貢献し、ビジネスやソリューションを次々と生み出して、持続性ある社会を構築していくという、こういった理念で構築できるのではないかと考えております。

4 ページ目「A I -Readyな社会」をご覧ください。

1 ポツの下から 2 行目ぐらいに書いておりますが、我が国としては Society 5.0 を実現するために、A I の実装を前提として社会構造をデザインし直して、「A I -Readyな社会」を世界に先駆けて構築する必要があるということ、「A I -Readyな社会」として 5 つ程度の要素を挙げて、それぞれ A I -Ready になっている必要があるといったことではいかがでしょうか。

人に関しては、A I を使いこなせるということ、A I のメリット・デメリット、データにバイアスがあるという特性もきちんと認識できているということ、創造性・生産性の高い仕事に従事できている、多様な人々が多様な夢やアイデアを実現できている、といったことを人の要素として挙げております。

社会システムとしては、医療、金融、保険、交通、エネルギー、さまざまな社会システムがありますが、これらが A I の進化に応じて柔軟に対応できるといったことが「A I -Readyな社会」ではないかと。具体的には、それぞれの社会システムについて拡張性や相互接続性を備えるような柔軟なアーキテクチャーが設計されて、それに基づくデータ利活用基盤が整備されている、と少々技術的な記述を置いてみましたが、単に柔軟と言うと意味が伝わらないと思いますので、少し具体的な内容を置いているところでございます。

産業構造も柔軟でなければならないと。企業における労働、雇用、それから創業環境が柔軟で、かつ国際的に開かれたものであるということ。

4 番目はイノベーションシステムということで、いわゆる研究分野です。大学、研究機関、

企業、一般の方々も含めて、イノベーションを次々とつくり出すことができるような環境であると。具体的には、リアル空間も含めたデータがAI解析可能なレベルで利用できるという状況と考えております。

5番目、ガバナンスということで、プライバシー、セキュリティ、公正競争確保といったAI・データに対するルール、制度の整備と、政府によるガバナンスが確立されている社会。社会的弱者を生み出さないようなルールや制度が整備できている、政府自身もAI-Readyになっているといったところをガバナンスとし、以上を「AI-Readyな社会」の5つの構成要素と考えております。

続きまして、AI原則です。前回、北野副議長からもご提案がありましたが、このAI原則の2つ目のポツにございますとおり、このAI原則を、「社会（特に、国などの立法・行政機関）が留意すべき」ようなAI社会原則と、AIの研究開発と社会実装に従事する開発・事業者側（がわ）が留意すべきAI開発利用原則といった体系化で整理してはどうかというご意見も踏まえて少し整理をし直しております。4.1が社会原則で、4.2が開発利用原則としております。

AI社会原則としては、1つ目は人間中心の原則として、人間の基本的な人権を侵すものであってはならない、人間がAIに支配されるのではなくて、AIの推論をうのみにしたり、悪用する人に操作されたりといったことがないように、リテラシーなどの教育をしっかりと整備することが望ましいということ。

要素としては、AIは人間の労働を代替するのみならず、人間の能力や創造性を拡充する、AIも人間から学習することで相互に高度化できるといったものであるということ、これまでの議論にもありましたが、AIの提案・推論に対して自らが判断を行うという、人間の自己決定権をしっかりと確認すべきであるということをございます。

2つ目、教育・リテラシーの原則は、AIリテラシーやスキルを身につける機会が平等に提供されなければならない、AI、数理、データサイエンスの素養を身につけられる教育システムが必要、AI・データの特性を認識できるような力が必要であると。さらに、そういった理数系のみだけではなく、創造性とか推論の構築、課題解決能力といったところの養成も重要であるということかと思えます。

6ページ目をご覧ください。プライバシー確保の原則として、AIを前提とした社会においては、個人の行動に関するデータから、政治的な立場、経済状況、趣味・嗜好などが高精度に推定できることがあり、単なる個人データを扱う以上に慎重さが求められる、プライバシーの

確保がしっかりされなければならないということでございます。3つブレットを準備してございます。

4つ目、セキュリティ確保の原則でございますが、さまざまなシステムの制御をAIが担う社会においては、システムが攻撃されるとこれまで以上に影響が大きいので、社会のレジリエンスを確保して、安全・安全で持続的な社会を構築するという観点からセキュリティの確保が重要であるということで、そのための原則として、研究開発、社会実装、サイバーセキュリティに関する研究開発や社会実装なども含めて、取組をしっかりとしていかなければならないというものでございます。

5つ目は公正競争確保の原則で、最近、新聞記事等にもあるように、欧州で大手のプラットフォームに対する規制の議論もありますが、支配的な地位を利用した不公正な競争や主権の侵害が行われる社会であってはならないということ。富や社会に対する影響力が一部のステークホルダーに過剰に偏ってはいけないということ。

7ページ目をご覧ください。FAT、公平性、説明責任、透明性の原則ということで、これは、人々が不当な差別を受けたり、人間の尊厳に照らして不当な扱いを受けたりすることがないように、公平性・透明性のある意思決定、その過程に対する説明責任が可能な限り確保される社会である必要があるということです。

7番目はイノベーションの原則ということで、イノベーションが次から次へとAIで生み出されていく必要があるという観点から、国境、産学官、ジェンダー、マイノリティといった垣根を越えてイノベーションが継続的に生み出されていく、そのための国際化・多様化、産学官連携がしっかりできなければならないということ。人材の移動も流動性がなければならない。AI工学の確立がされなければならない。それから、データも国境を越えて相互利用できる。プライバシーやセキュリティの確保が前提とはなりますが、国境を越えて相互利用できる。それから、AIの社会実装の阻害要因となっている規制はしっかり改革されなければならないといった原則でございます。

4.2、開発利用原則でございますが、AIシステムを開発し、サービス・製品を展開しようとする関係者は、上記の理念や社会原則を踏まえたAI開発利用原則を定めて遵守すべきであるとしています。

AI開発や利用に関する原則は、現在、多くの国、団体、企業において議論されておりますので、国際的なコンセンサスをしっかり醸成しながら、非規制的で非拘束的な枠組みとして国際的に共有することが必要であるということで、現在のAI開発利用原則の例をここに記載し

ております。

このあたりは、先ほど来話題になっております、EUにおけるAI原則の動向なども見ながら、開発利用原則をどの程度この場で議論するのがいいのかといったところも、少し今後検討できればと考えております。

最後に、8ページ目でございます。

「おわりに」ということで、前回会合で結びが国際的な情報発信だけで締まっているのが味気ないのではないかというコメントも頂きまして、今後内容を詰めていく必要があると思いますが、今の要素として3つほど挙げさせていただいております、今後、「AI-Readyな社会」を世界に先駆けて構築していくため、本原則は、政府、関係機関、団体等で共有・認識していくということ、国際的な議論の場において世界各国と連携・共有することで、国際的なコンセンサスの形成を目指すこと、本原則はAI技術の進化や社会の変化に応じて柔軟に見直すといったところを記載しております。

机上資料1については以上でございます。

#### 【須藤議長】

どうもありがとうございます。

机上資料2が皆様のお手元にあると思います。議論のための参考として、北野副議長に作成していただいたものです。簡単に北野副議長からご説明をお願いします。

#### 【北野副議長】

これは、今いろいろなところで議論されているAI原則や政策パッケージの、我々の俯瞰（ふかん）図になるのではないかと思います。

現在、内閣府・官邸の統合イノベーション戦略推進会議の中でAI戦略が議論されていまして、これが社会原則、AI原則であるとか、教育も含めた全体のパッケージになっているわけですが、その中で、本会議では人間中心AI原則というのが議論されていて、そこで開発原則と社会原則というのが議論されているというところであります。同時に、先ほど申し上げましたように、昨日公表された経団連の未来社会協創会議のSociety 5.0提言書というのがございまして、経団連のこれからの一つのフレームワークになるものでございますが、その下に、私が議長を仰せつかっているAI活用原則タスクフォースがございまして、このタスクフォースから経団連としてのAIの原則の提言書が、年内または年明けぐらいに出ることになります。

これは経済界、事業者側（がわ）としての立場でございますので、開発原則はその中に当然入るわけですが、同時に経営側（がわ）として、日本の企業がどうやってA I-Readyに変革していくかということに対する企業としての原則——中身に関しては、ここに書いてあるものは、議論をこれからしていきますので、変わるという前提でご覧いただきたいのですが——を書くこととなります。

その中で、さらに政府として、または世の中としては、こうあるべきではないかというリファラーはするとは思いますが、経団連の方のミッションとしては、基本的には事業者側（がわ）と、それを経営として企業はどう変革するべきかといった内容になると思いますので、おそらく、全体像が出たときには、この会議のガバメント、インターガバメントレベル、それと開発原則、ここは経団連とオーバーラップして、経団連側（がわ）は経営側（がわ）の企業変革に対するA I-Ready化ガイドラインのようなものが出ることになって、これらが連動するという形になるのではないかと想定しております。

以上でございます。

#### 【須藤議長】

どうもありがとうございます。非常に鳥瞰図としてよくできておりまして、参考にして、また議論していただければと思います。ありがとうございます。

それでは、先ほど新田参事官からご説明いただきました再構成案について、構成員の皆様からご意見、ご質問を頂きたいと思っております。論点が拡散しないようにするために、まずは理念のところ、次に「A I-Readyな社会」、それから原則について、と3つに分けてそれぞれご議論、ご意見、ご質問を頂くようにしたいと思います。最後に、残り時間を使って、枠を設けず、全体の構成についてご議論を頂くことにしたいと思います。

執筆構成、執筆していただく方々については、私と北野議長と事務局でいろいろ議論していて、ここら辺はこの方々に書いていただきたいというのは議論してきたのですが、まだ本日は出せません。本日のご議論でまた構成が変わる可能性がありますので、それを踏まえてお願いを申し上げるということにしたいと思います。

それでは、まず理念のところについて、ご質問、ご意見を伺いたいと思っております。

#### 【中川構成員】

理化学研究所の中川です。理念のところ、非常によく書けていると思って感心しておりますが、①、②、③で、ディグニティ、人間の尊厳が尊重される世界というのがあります。実際、人間が尊重されるのはそのとおりなのですが、これに関連して、A Iも何らかの基本権を持つという考え方がEUなどでも出てきておりまして、人格権というわけにはいかないのですが、何らかの権利をA Iに与えた方が、法制度的に見てもすっきりした形になる。

例えば著作権も、何らかのA Iの権利の持たないと、うまく使えずに、出版会社の人だけが得をするとか、著作隣接権の争奪戦になるような状況になる。これは余りうれしくないのも、何らかの方向性をA Iに対する権利の与え方として考えてもいい時期に来ているのではないかと、EUではそういう少々先走りの意見が夏頃出ていましたが、世界中でそういう雰囲気が出つつあるということをおもいましたので参考まで。

**【須藤議長】**

ありがとうございます。

ほか、いかがでしょうか。

**【安宅構成員】**

私も、この理念、とてもよくできていると思うのですが、2番目のダイバーシティ・アンド・インクルージョンのところですが、多様なものをのみ込めるようになるということはもちろんです。これに加えて、多様から価値を生むという感じの部分も一言入れておけるといいかなと思いました。

以上です。

**【須藤議長】**

ありがとうございます。

ほか、いかがでしょうか。

**【丸山構成員】**

細かいことですが、2番のダイバーシティとインクルージョンの、「インクルージョン」の部分が下のパラグラフにないので、何か入れられたらいいかなという気がします。

**【須藤議長】**

ありがとうございます。

ほか、いかがでしょうか。

羽鳥構成員。

**【羽鳥構成員】**

日本医師会、羽鳥です。

本当に僕もよく書かれていると思いますが、例えば、これは、最後の方で議論した方がいいのかもしれませんが、経団連の方から、医療の情報バンク構想とかもお持ちだということですが、EHR/PHRを個人で持たせる、あるいは医療機関で持たせるのではなくて情報銀行に人間の検査データを持たせるような仕組みをつくらうという発想もあるようですが、フェイスブックなどの情報漏洩や、トランプ大統領のSNSを利用した論理より感情に訴える仕組みで集票するやり方を見ていると、倫理的、法的な縛りがあってもいいのかなという気がします。

以上です。

**【須藤議長】**

ありがとうございます。

北野副議長。

**【北野副議長】**

今の羽鳥構成員のご意見は、実装的なところですので、理念というよりもおそらくAIの社会原則の中の制度設計のところ、プライバシーやリテラシーのところに書き込む形になるかと思います。

中川構成員のAIの権利は、おもしろい議論なのですが、AIに権利といってもおそらくそのオーナーが誰かいるとか、そういう問題もあるので、その辺はうまく議論をしながら、書き込めるかどうか状況を少し見定めていきたいと思います。

よろしくお願いします。

**【須藤議長】**

どうもありがとうございます。

羽鳥構成員のご意見について、私も北野副議長と同意で、原則のところのプライバシー等のところで明記すればと思います。

中川構成員、ただいまの北野副議長のご意見についていかがでしょうか。

**【中川構成員】**

そういう考え方が出つつあるというご紹介ですので、入れるかどうかはご議論いただければいいと思います。方向性として少し出てきたという紹介です。

**【須藤議長】**

ありがとうございます。

ほか、いかがでしょうか。

よろしいでしょうか。

後ほどまた振り返って、全体についてご意見、ご質問を伺いますので、その際にでもまた発言いただければと思います。

続きまして、3の「AI-Readyな社会」についてご意見、ご質問を承りたいと思います。いかがでしょうか。

安宅構成員。

**【安宅構成員】**

安宅です。

これもすごくよくできていると思いますが、2点、可能であれば回答いただけるといいなと思います。

一つは「人」のところで、これは統合イノベーション戦略会議の方にもうかなり書かれているのですが、理文や専門といた枠に捉われず、データやAIの基礎教養、実装に向けた設計とかデザイン力を持つ人材が十分に存在している系の話を入れられるといいなと思ったのが一つ。

二つ目は「イノベーションシステム」、これもすごくよく書かれているのですが、「リアル空間も含めたあらゆるデータが安全に」、のこの「安全に」の前に「新鮮かつ」という言葉を入れていただけるといいかなと、データというのは鮮度で価値が決まるものですから、これは

非常に重要で、半年後に使えるではしようがないということで、2秒後には、できたら0.5秒後に使いたい、そういう感じです。

**【須藤議長】**

どうもありがとうございます。

「新鮮な」と、これもうまく使えるように何とか入れ込みたいと思います。ありがとうございます。

ほかはいかがでしょうか。

堀構成員。

**【堀構成員】**

東京大学の堀でございます。

「イノベーションシステム」で、細かい文言なのですが、イノベーション論の方々は、イノベーションは創り出すものではなくて生まれるものだとおっしゃる方も多いので、ここは「イノベーションが次々に生まれる環境ができていく」の方が、つくろうと思ってつくれるものじゃなくていろいろな相互作用から生まれますので、文言としてはその方がいいかと思いますが。

**【須藤議長】**

ありがとうございます。

私も賛成です。いろいろな力あるいは人や物が加わって生成されるという感じですので、「(イノベーション)を」じゃなくて「(イノベーション)が」ですよね。「が」がいいと思います。ありがとうございます。

**【中川構成員】**

今の安宅構成員の意見はそのとおりだと思いますが、それを言うとおそらく、データのソースが誰かというアイデンティフィケーションの問題が必ず入ってきてしまう。特に個人データの場合は、本当はネットワーク越しでアイデンティフィケーションをしなくてはいけないという問題が入ってくるのですが、その辺が正しくないデータをベースにして進めるということ

がしばしば行われる。認証された正しいIDも含めて入っていた方が、新鮮さと同時にAI-Readyなデータではないかと思います。

**【須藤議長】**

ありがとうございます。重要な点だと思います。

ほかございますでしょうか。

どうぞ。

**【浦川構成員】**

損保ジャパン、浦川です。

このAI-Readyな社会の部分で語るべきかどうかなのですが、前段の基本理念の三つのポイント、三つの価値を尊重するということが、具体的に次のパラグラフ並びにこの次のAI原則にどのようにつながるのか、前段の三つの価値が、AI-Readyの五つのポイントや原則にどうつながっていくのかということ、どこかでつながりを図示するか何かするといいいのかなと、例えば前段の2点目のダイバーシティは、前回も申し上げたとおり若干AIと照らすと分かりづらいので、結果的にはこれは4点目のイノベーションシステムのような格好でダイバーシティが実現されるとか、分かりやすさを追求する工夫をしてはどうかとご提言申し上げます。

**【須藤議長】**

ありがとうございます。

これについては何かほかにご意見がありますでしょうか。

福岡構成員。

**【福岡構成員】**

現状はAI-Readyになっている必要があるものについて5つ挙げられていますが、先ほどのご議論をお伺いして、データ自体もAI-Readyになっていることが重要なことだと思っております。実際使うときにデータがAI-Readyになっていないために非常に苦勞するという話はよく聞いておりますので、ご検討いただければと思います。

**【須藤議長】**

ありがとうございます。重要な点です。

ちなみに先々週、エストニアのタリンに行ってきましたが、X-Roadによって十数年前にデータは全部タグが統一されていて、もう既にA Iで活用する前に持ってきています。とっくにA I-Readyになっているよと言っていましたので、ここで「世界で我々が最も先駆的に」と書いてありますが、そんなこともないので、これは消しておいた方が、エストニアとかに「おまえ、うそ言うな」と言われますので、ちょっと書き方を変えた方がと思います。

ありがとうございます。ほかはいかがでしょうか。

どうぞお願いします。

#### 【樋口構成員】

私もよく書けているんじゃないかと思いますが、(1)の「人」のところで「データにバイアスが含まれる可能性がある」と書いてあるのですが、違和感があって、「データには多かれ少なかれ必ずバイアスが含まれる」と言う方がいいと思います。バイアスの存在を可能性と言うのは少々問題があると思いました。

#### 【須藤議長】

ありがとうございます。これも統計処理上極めて重要な問題で、統計学はそこら辺を検証し、それからどのぐらい偏奇があるかというようなことを明らかにするというのが、重要な作業になっております。樋口構成員のご見解はもっともですので、ここもそういうきちんとした記述にすべきだと思います。ありがとうございます。

ほかはいかがでしょうか。

お願いいたします。

#### 【高原構成員】

トヨタ自動車の高原です。

中国の北京近くの雄安という都市では、新たにA Iで駆動した新たな交通体系と新しい空間秩序をつくる都市をこれからつくっていくということが、多く報道されていますが、そういう中でこの社会システムのところに、新たな空間秩序を形成するような「都市」などの言葉があるといいのではないかと思います。

**【須藤議長】**

ありがとうございます。重要だと思います。空間として都市、都市じゃなくてもいいのですが、その空間について具体的なものを書くということは重要だと思いますので、ありがとうございます。

ほかはいかがでしょうか。

お願いいたします。

**【新居構成員】**

(1) と (2) のところで「多様な人々が多様な夢やアイデアを実現できている」ということが2回登場していて、これだけ短い文章の中で2回出すということはかなり強いメッセージになってくると思うので、果たして同じ表現を2回使うべきなのかということは、今後細かい文章修正はあると思いますが、考えた方がいいかなというところと、あとは社会実装に関しては冒頭の部分でも触れられていると思いますが、OECDでも信頼構築と社会実装の二つをキーワードとしている中で、信頼構築の議論はこれまでもあったが、入れていないというところも気になりましたので、その辺も入れていくと、より踏襲している感じが出るのかなと感じました。

以上です。

**【須藤議長】**

ありがとうございます。私も信頼について記述を入れた方がいいと思います。何とか入れようと思います。それから重複があるところは、変にひっかからないように何とか考えるべきだと思います。ありがとうございます。

ほかはいかがでしょうか。

丸山構成員。

**【丸山構成員】**

2点ありまして、2の「社会システム」のところでは、効率化だけではなく多様な目的関数を持つようなシステムみたいな言い方が入るといいかなという気がいたします、ITを入れると直ぐコスト削減となる人がいますので。

それから5の「ガバナンス」ですが、ガバナンスが確立されていると同時に、この領域は技

術がすごく速く進むので、ガバナンスがアジャイルに動いていけるようなガバナンスであってほしいというような気がいたします。

**【須藤議長】**

ありがとうございます。重要だと思います。特に2番目のガバナンスのところのアジャイルなガバナンスというのは、これはかなりグローバル、国際会議でもインパクトのある言い方になる、いい提案だと思います。ありがとうございます。二つとも分かりました。これも事務局でまとめます。ありがとうございます。

羽鳥構成員。

**【羽鳥構成員】**

5番の「ガバナンス」に関してですが、いわゆる真実のデータとにせのデータ、先ほどのフェイクニュースとも関係するのですが、例えば医療の現場でも起こりうることで、自分の知らない病気の概念を知りたいと思って、例えばグーグルで検索すると、情報を出したい診療所、病院、OTCの製薬企業は食品のメーカーがこういうのに効きますよ、というような自己の宣伝がSEOといわれるアクセス数に応じたものがトップに来るような仕組みになっています。ヤフーは改善しつつありますが、そういう真のデータとにせのデータが分かるような仕組みにできないか。本来は良貨が悪貨を駆逐してくれればいいのですが、今の世の中は良貨が悪貨に駆逐されてしまうような状態なので、それを何か工夫してほしいなと思います。

例えば米国の4分の1の人口の福音教会の人が進化論を否定して神が人をつくったと、それは4分の1の米国人は信じていると、そういうことがあるのは誰が見てもおかしいと思いますが、

そういうのをどこかで何か書き込めないと、常にニュースを出したい人、情報を出したい人の一人勝ちになってしまって真の情報が隠れてしまうのではないかと思うので、何かどこかでうまく書き込めればお願いしたいと思います。

**【須藤議長】**

ありがとうございます。これは恐らく先ほどご意見いただいた「信用」とか「信頼性のある」

という言葉が重要になりますので、そういう形でこの3番目のAI-Readyな社会のところは書き込ませていただけたらと思いますが、この取りまとめがもうちょっと全体的にボリュームが大きくなる場合には、今、羽鳥構成員がおっしゃっていただいたような具体的なことも、触れるのかなと思います。ありがとうございます。

どうぞ。

**【北野副議長】**

今のところは、もしかしたら開発者原則なのかなという気もします。外から、またガバメントレベルからどういう情報を出す、出さないというと、これは検閲になる可能性もあり、そこは少し微妙なバランスになってくるので、重要なことであることは間違いありませんが、それをどこで書き込むかというのは、うまい書き込み方をする必要があるかなとは思っています。

**【須藤議長】**

ありがとうございます。

どうぞ。

**【木俣構成員】**

今までの議論を聞いていて、データの信頼性といったところという意味では、トレーサビリティというようなキーワードも入れてもいいのではないかなと思いましたが、いかがでしょうか。

**【須藤議長】**

ありがとうございます。私自身はトレーサビリティという概念は入れるべきと思いますが、使う場合は定義をある程度使った方がいいですね。その上で社会の信頼性とかデータの信頼性とかを包含できるような使い方ができると思います。ありがとうございます。

ほかはいかがでしょう。

どうぞ。

**【関口構成員】**

一点目は「ガバナンス」のところですが、先ほど、丸山構成員がおっしゃられたようなアジャイルに回していくというのは大賛成で、私も同じように思っています。つけ加えるとすれば、政府によるガバナンスがここには明記されているのですが、データを扱うコミュニティーなり、一つの試行的なグループが、ある業種、業態のところでは先行して回していけるような方法があっただけではないかとも思いました。

二点目は「イノベーションシステム」に関してなのですが、二つ目の丸のところ、データがレディーになっていなさいということが書いてあります。データがレディーになっているだけでアクセスされないとデータの持ち腐れにもなります。冒頭にもありますようにアクセスのネットワークがきちんと整備されていること、ここは超高速広帯域通信網と、5Gを含めた高速なネットワークというのがあることと、データが光速に処理できるような環境が整っているということが、全てパッケージになって提供されることを忘れてはいけないと思います。

**【須藤議長】**

ありがとうございます。

北野副議長。

**【北野副議長】**

3のところですが、実はこの書き方について事務局と議論をし、次の4でAI原則があつて、4.1が社会原則で4.2が開発原則とあるので、ここ方向性を示しコンパクトな感じでまとめています。ですので、ご指摘のところは全て重要なことは理解しているのですが、実は余りここには書き込んでいなかったのです。ただ、今の議論をお伺いすると、AI-Readyな社会としてあるべき姿は、かなりきちんとここに書き込むという方向にした方がいいのかなという気もしてきていて、こういうのがAI-Readyな社会とあつて、それを実現する枠組みとしては社会原則をこうやる、開発原則はこうであるとなると、この3のところのボリュームは、今想定している抽象的な方向性というよりも、もう少ししっかり書き込むという方向になるのかもしれない。どちらにするかによって、構造はそんなに変わらないと思いますが、構成上大分ウェイトが変わって、ここにどのくらい書き込むかの重さは変わってくるかなと感じています。これは後でおさらいのときに議論したいと思います。

**【須藤議長】**

ありがとうございます。重要なところで、ここはまた議論をしたいと思います。

平野構成員。

**【平野構成員】**

今の副議長の話が4のところに触れられていたので、4のところでお話ししようかと思っていたのですが、4.1と2と二つ設けられて、その2の方の開発利活用というところは、ご承知のようにAIネットワーク社会推進会議で須藤議長が座長になられて、1年、2年かけて相当回数議論しているところですので、ここで短期間ではちょっと深掘りができないなと思います。ですから私のアイデアなのですが、ここは余り、今、深く書こうかなというご提案はあったのですが、特に4.2はむしろ深く書かないで、そういうものがありますよとリファーする方がいいのかなと、4.1は今回すごく新しいので、そこに選択と集中というか、集中した方がいいのかなというのがご提案でございます。

**【須藤議長】**

ありがとうございます。この後北野副議長にもご意見いただきますが、私も北野副議長も事前の議論では事務局も、その方向でいこうということを言っています。

北野副議長。

**【北野副議長】**

4.2は正にいろいろなところで議論されていてそれなりに相場観ができていますので、ここをもう一回議論し直すのは時間も意味もないと思います。ですので、こういう議論がありますよとリファーして再掲するような形で、ここでの書き方も項目を並べるだけで非常にリーンなものになっていますが、おそらくそうなると思います。4.1が今回の非常にオリジナルなところになりますが、ただ、4.1の社会原則だけ出すと、4.2の開発原則があつてこれは開発原則じゃなくて、いろいろな原則の関係性が分からなくなるのはまずいと思いますので、全体の体系は明確にしておいた方がいいだろうと。それに基づいて、なぜこの社会原則を我々が特に打ち出していて、開発原則を割と相場観があるところの上でやっているかということ、3として我々が目指すべきAI-Readyな社会があるからこういうのが必要だ、ということにおそらく

なるのかと。そうすると3のところのボリュームをどのくらい重く濃くするかどうかというのが、一つの思案のしどころではないかと思っております。

#### 【須藤議長】

どうもありがとうございます。そのとおりです。それで3はできるだけスリム、リーン、ぜい肉をそぎ落として、こういう社会が求められるよと、では、こういうAI-Readyにするためにはどういう原則が必要かというところで、この社会原則のところをかなり記述していこうというのが、ここに提出する前の北野副議長、須藤、それから事務局の意見でございます。

したがって、もう既に4のところの原則に議論が移ってきておりますので、続きまして4を議論していただいて、これまで頂いたご意見は4に吸収できるものは4に持っていきたいと思います。もちろん概念としてこういうコンセプトを入れておくことが、AI-Readyというところで必要であれば入れますが、基本的には議論の充実はこの4のAI原則、5ページ以下で展開させていただければと思います。そのような観点からご議論いただければと思います。どうぞ。

#### 【中川構成員】

4も大変まとまってよく書けていて頭にはすっきり入るのですが、そうすると逆に欠けているかなと思う点が出てきます。AIが何をしているか分からないというブラックボックス化というのがよく議論されていて、それがために信用を失い、なかなか企業での導入が難しいというような話もよく聞きます。それに対して対策をきちんと立てないとみんなが使ってくれない、安心して使えないという状況になるのではないかということで、そのことについてもう少し書き込んだ方がいいかなと。具体的に言いますと、「人間中心」の最後のブレットのところ、AIの提案・推論に対して自ら判断を行いとありますが、提案や推論に対し自ら判断を行える人というのは相当リテラシーの高い人であって、これは普通のリテラシーの一般人が判断できるためには、逆にどういった情報をAI側（がわ）から提供してあげればいいのか、一般の人に理解可能な形での提供というのは非常に重要な技術的ポイントないしは、社会に対して影響を与えるポイントになってくるので、もう少し書き込んだ方がいいのではないかと思います。

それに関連して、非常に複雑なデータ環境になっているので、プライバシーセキュリティ等にも絡みますが、外部の複雑なデータ環境に人間の知能だけで対応できない状況にだんだん

りつつあります。ですからA Iのエージェント、個人の代理人のようなA Iとのエージェントを人間と外界の間に入れて、そして外側（がわ）とのやりとりがある。個人情報保護なども明らかにそういう側面があるのですが、私がI E E EのE A Dにコミットしているのですが、A Iエージェントのような、生まれる前から死ぬまで代理でしてくれると、そういったようなタイプのA Iが必要なのではないかというのがE A Dのバージョン3で出てきますが、そのような考え方を入れていく。すると、人間と対抗するA Iではなくて人間の代理をしてくれるようなA Iという概念が、もう少し表に出てきてもいいのではないかと、そういう方向での議論も進んでいるということの一つ申し上げたかったということです。

最後にもう一点、先ほど信頼という言葉がありました。一般の人に理解可能であるということは、必ずしもトランスペアレンシーがあるとかトレーサビリティがあるとかということだけで実現できるのではなくて、ある程度トラストしてもらおうという部分がどうしても必要になってくる。トラストというのはある意味では思考停止をするということと等価なので、どういう状況でトラストできるんだらうと、これは大屋構成員がお詳しいと思うのですが、その辺りは実際のA Iが社会に組み込まれて安心して使ってもらうときの非常に重要な要素になるかなと思いましたので、一言。

**【須藤議長】**

ありがとうございます。前半のところの人間の判断とそうじゃないところのエージェントについても議論したいのですが、2番目の信頼、トラストというところ、連動するのですが、大屋構成員の名前が出ましたので大屋構成員、ご発言いただけますか。

**【大屋構成員】** ありがとうございます。その点も含めてちょっと申し上げさせていただきたいと思います。

トラストの話から始めますが、1番の「人間中心の原則」の、先ほど中川構成員がご指摘されたところの書き方がうまく考えられていると、つまりここでは、あくまでA Iは提案・推論をするので、人間が判断をするのだという扱いになっている。それと平仄を合わせる形で6番の「公平性、説明責任、透明性」のところでは人間が主語になっているわけですね。つまり第2点で「データの取得方法や使用方法を含めたA Iシステムの動作について、可能な限り説明が得られなければならない」となっていて、人間の側（がわ）から見たときに説明と感じられ

るものが出てくるかという話になってくるわけです。

ただ、この意味で要するに人間が決定者であるので、その人間が最終的には責任を負うのであると明確になっているところが、筋としては非常にいいと思うのですが、一方で今、中川構成員がおっしゃったように、判断を行いというのが、判断はできるのかしらというところが問題になってくるわけですね。ここはおそらく現実的にはということになるのですが、リテラシーの高い人は実際に判断をするわけですね。例えば病理診断のAIがレントゲンを読んだものを専門の病理医が見るというケースでは、念のためと思って見たが、大丈夫だねとか、ちょっとおかしいなと思って判断をするわけでしょう。ところで、リテラシーの低い人はそれができないので、何となく正しそうだけど、いいかなという「ゴー」か「ノー・ゴー」かの判断をしていると、これは判断というよりはどちらかというところと決断に近いわけです。ただ、いずれの場合にせよ、そこでゴーだという決断をしたからにはその責任を負ってもらう必要があるのである、AIがその責任を負えない、ということを明確にすると、責任分界点としては切り分けができるのかなという気がいたします。

仮に一応そういうたてつけであるということ的前提とすると、関連して一つはそこで「人間の自己決定権」と言うところとちょっとおかしくて、「人間の決定権」とすべきではないかということと、6の「説明責任」の先ほど申し上げたところで、一つは、「可能な限り」と言ってしまっているのかどうかというのはちょっと微妙なところがあると思います。人間を主語にした場合には、どちらかというところとAIシステムの動作を理解するための材料が、例えば可能な限りか適切な水準で保存・提供されなければならないという形で言っておくといいのかなという気がいたします。これは保存と提供を分けているのは、場合によって、例えばログが膨大であるとか、あるいはそこにもう何らかの意味でトレードシークレットがかかるというようなケースにおいて、一般的に提供はできないが、例えば司法機関の捜査が入った場合には提供できるようにしておいた方がいいというようなこともあり得るだろうということで、そのような形にしてはどうかと思っております。

あと中川構成員のおっしゃったところではトラストですね。トラストについてはいろいろなタイプのトラストがあるというのか、これは政治学の分野では大体、権威の問題として分析をするのですが、権威が成立するためには何があるかということ、カリスマ的な権威みたいなものを抜きにしておくと、合理的であるか、要するに責任が担保されるか、要するに動作が判明するので、なるほどこういう理屈で動くなら大丈夫だという形で信頼する、こういう理由で判断されるなら大丈夫だということで信頼するというのが一つ、それからエキスパートイズが、例

例えば医者の場合は典型的だと思いますが、エキスパートイズが保障されており、あと医者のコード・オブ・コンダクトがかかっていると、医者というのは医療倫理がかかっている、悪いことはしないはずであるという形で信頼が生じるというケースもあると思います。なので、何かあった場合には医療事故に関する責任をとる制度というのがあるはずであるということで信頼が生じる。

そうすると、以前にアカウントビリティーの説明のところでもちょっとお話しさせていただいたと思いますが、中身はこうなので、それを理解した上であなた自身でご決断くださいというのが一つ目のアカウントビリティーのとり方であり、中身については詳しく言えないが、何かあったときには私がアカウント・フォーシマスと、つまり動作の結果については責任を負いますというのがもう一つの責任のとり方であると、この二つに対応するような形できちんと場合分けをして書けるといいのかなとは思いました。

ついでに2点ほど申し上げていいですか。

一つは4番のところですが、ここではセキュリティ確保ということで悪意あるサイバー攻撃、かなり大規模なものを想定して書かれているのですが、先ほど、羽鳥構成員がご指摘されたようなフェイクニュースであるとか、あるいは意図的な誇大広告に関わるような宣伝というもの、データインテグリティーとか信頼性を毀損する行為ということで、ここに入れてもいいのかなという気はいたします。

ただ、ここではサイバー攻撃、データインテグリティーについては、これは当然だと思いますが、かなり強い言葉で批判をしている。これに対し、例えば企業の一般の経済活動でもそうかもしれませんし、通常の個人の情報発信の場合には、意図せずしてフェイクになってしまうということは当然あり得るわけであって、意図的にデータの信頼性であるとか情報を混乱させるような利用をしてはいかんというぐらいの書き方に、とどめるといいのかなとは思いました。

3つ目は、これはごく単純な話ですが、1番の「人間中心の原則」の1行目に、「自由・尊厳・平等など、人間の基本的人権を侵すものであってはならない」と書いてありまして、「人間の基本的人権」というのは疊語ですというのが一つと、ここはできれば「人間の基本的価値」などにさせていただきたい。というのは、人権として実定化されたものだけがここで問題になるわけではないでしょうと思うところであります。

以上です。

【須藤議長】

ありがとうございます。とても説得的で興味深く伺いました。

お願いいたします。

**【永沼構成員代理（江川）】**

今の大屋構成員の説明の中で、説明ということと、それから中で何が起きているのかというご議論がありましたが、標準化の世界では正にそれが今いろいろと議論されておりまして、人間が聞いて分かる説明というのは、エクスプレーナビリティ、説明性という言葉で議論する方向でほぼいろいろな標準化団体がまとまりつつあります。

そして専門家が見たら中はどうなって、本当のところ何が起きていたのか、例えばログデータ、これは人間がただ見るだけではなかなか理解できない代物ですが、でも、専門家がきちっと解析すると本当のところ何が起きていたのかが分かると、これはトランスペアレンシーという言葉で呼ぼうということで、幾つかの標準化団体が大体合意に達しつつあります。

したがって、そういう用語を使っただけだとよろしいかなと思いますが、その場合ここで述べている透明性の原則というのは、社会の対話の方に話が振れていて、この6のところ、これはこれで重要な透明性だと思いますが、ちょっと議論が混乱するので、そこの書き方を注意しないといけないかなと思う次第です。

以上です。

**【須藤議長】**

ありがとうございます。専門家として了解可能なというところも含めて、うまく一般向けのものと、それから専門家で相互に了解可能な情報あるいは定義あるいはデータということ、もう少し議論はしておいた方がいいと思います。ありがとうございます。

ほかはいかがでしょうか。

堀構成員。

**【堀構成員】**

東大の堀です。

A I 社会原則の議論は国内にとどまらない側面があるのは明らかだと思いますので、机上資料2の北野副議長の資料にあります一番下の国際的枠組みの構築というのを、是非独立した項目として入れていただいた方がよろしいのではないかと思います。

【須藤議長】

ありがとうございます。そのとおりでしょう。それは強調した方がいいと思います。ありがとうございます。

【北野副議長】

先ほどの大屋構成員の質問ですが、「人間中心の原則」のところで書いていて、少し整理をされていて一番悩んだ部分というのがあって、AIが提案とか推論を行って人間が判断するという書き方をしているのですが、そういったときにきちんと判断ができるかどうかという先生の議論がまず一つあって、それは正にそうであると思います。もう一つあるのは、例えば自動走行のようになったときには、自動走行システムにある意味デリゲート、権限委譲して自動的に連続的に判断、実行をやっているわけですね。人間が判断するという原則をそこに置いた場合には、アクセルを踏め、ブレーキだ、右だ、左だ、を言いながら人間が操作を行うというシステムになるわけですよ。これは逆に危なくてしょうがない。実際に自動走行をやるときには完全に権限委譲して、しかも命に関わることを権限委譲しているわけですね。もちろんそこから離脱する権利というのは当然ありますが、そのときに人間が最終的に判断するべきであると書き込んでしまった場合、権限委譲したときは権限委譲するという判断をしているからいいとするのか、そこの書きぶりが少し悩ましい。

それともう一つは、AIが何でこれだけ重要になっていくかということ、かなり多くの場合きちんとしたトレーニングデータが確保できて実用になるとときには、人間よりも大方いい判断をAIの方がしているからこれを使おうとなります。このときに、最終的に人間が判断するのは、人間の尊厳のために問題を引き起こすことを許容するというジレンマに陥る可能性があって、そこをどう書くかがすごく難しいのです。ここはきちんと議論しておかないといけないかと思います。そこはすごいジレンマなのでご意見を頂ければと思います。

【中川構成員】

今の北野副議長のご意見は正にそのとおりで、ですからディグニティのところである種の権利、人格とまでいかないにしても、AIはある種の権利と責任を負うというような位置づけに、もうそろそろしなくてはいけないのではないかと。今のレベル5の自動走行は、もうそういうのはだんだん出てきているのですが、では誰が責任をとるんだろうということを含めると、どう

しても何らかのA Iの責任のとり方あるいは権利、基本権みたいなところに、踏み込まざるを得ない状況になっているのではないかと、今のお話をお聞きして議論していただきたいと思います。

【須藤議長】

それでいくと、中川構成員、その推論は興味深いのですが、法人格を組織に与えたような擬制的な人格というか、そういうものをA Iに与えるという議論も、今、国際会議でなされていますが、それにつながる議論に持っていくべきだということですか。

【中川構成員】

それは一つの方向性としてあると思います。

【須藤議長】

ありがとうございます。

どうぞ。

【大屋構成員】 すみません、今の点だけ簡単に補足させていただきます。

これは二つの答え方がおそらくあって、現在の法体系の枠組みの成り立ちから見直すのであるならば、中川構成員がおっしゃったように例えばA Iに法人格を認めることも、そんなにすごい修正ではない。A Iに法人格を認めて、A Iの責任を問うということはある選択です。というのは、責任は人格にしかかかりませんので、人間からA Iの方に責任を投げたいのであればどうしてもそこに人格がないといかんというのが原則になると思います。ただし、人格をつくるのはいいのですが、人格があっても損害に対して引き当てる財産がないと、社会的に損害を投げただけになってしまうので、裏打ちになる財産制度みたいなものをどうデザインするかということとセットでないと、事実上は無意味になると思います。

そこまでとりあえず踏み込まないのであれば、法の建前からいうと、今述べたように人格のある人間しか責任を背負えないので、A Iに人格を認めないとすれば人間が背負うしかありません。という場合は判断の概念を拡張して、要するに先ほどのデディケーションであるとか、あるいは利用する・しないという判断が、責任の根拠たり得るという構成にするしかありません。

例えば自動運転車、レベル5なんかの場合には、乗った人に乗るという決断をしたところが多  
少ある。もう一つは自動運転車の製造者か運行供用者に、そのようなシステムを利用に供した  
と、公衆の利用に提供したというゴーサインを出したという判断があり得ると、その両者が重  
畳して一定の責任が生じるという構成になろうかと思えます。

【中川構成員】

すみません、今の点について一つだけ、正に大屋構成員がおっしゃるとおりの方向で、二つ  
どちらをとるかを挙げなくてはいけないのですが、具体的にはどう人々を救うかという保険  
というのを無視することはできなくて、多くの場合は保険制度をどう構築するかということと  
絡んでくると思っています。ですから保険というものをA Iの利用の中で位置づけていくとい  
うのは、ここでは全然書かれていませんでしたが、もし実際に実用上で考えるのであれば余り  
無視できない部分かなと思っています。

【須藤議長】

ありがとうございます。

丸山構成員、それから平野構成員。

【丸山構成員】

大屋構成員の2番目の解釈に賛成でございまして、この文言を「A Iの提案・推論に対し」  
という非常にスペシフィックな言い方ではなくて、「A Iの利用に当たっては」ぐらいにすれ  
ば、それで済む話のような気もいたします。

【須藤議長】

ありがとうございます。

平野構成員。

【平野構成員】

文言ではなくて中川構成員のご質問のところ、自動運転についてはもう報告書までできて  
いるので、その保険の部分ですね、本日は国交省からも出席されていらっしゃる、もし時

間が許せばご発言を頂きたいなと思いますが、どうでしょうか。

**【須藤議長】**

国交省、用意できますか。ご発言できますか。

**【菊池専門官】**

関係部局とも相談の上、ご回答させていただければと思います。

**【須藤議長】**

報告書が出ているということですので、それに基づいてまた次回にでもちょっとお話しただければと思います。

樋口構成員。

**【樋口構成員】**

「教育・リテラシーの原則」のところの「このリテラシー教育は」から始まる場所なのですが、ここで2行目に先ほど、私が指摘させていただいた「バイアスが含まれる可能性がある」と、先ほどの論点プラスこの書きぶりについてちょっとコメントがあります。バイアスが必ずあるとともに、使い方によってバイアスを生じさせることも可能でありますので、その辺のことというのは実は（6）の公平性あるいは公正性というところにもつながってくるので、この書きぶり見ますと、最後に「創造性、推論の構築、課題解決能力」というように、クリエイティビティーに言及しているのですが、公平性や公正性を正しく理解するということが非常に重要かと思っておりますので、その辺りもこのリテラシーのところに書いておくといいのではないかとこのように思いました。

**【須藤議長】**

ありがとうございます。

なるほど、そのとおりかもしれません。重要なお意見だと思います。

ほかはいかがでしょうか。

**【福岡構成員】**

二つあります。まず、4番のセキュリティ確保の原則というところですが、通常セキュリティといった場合には、セーフティとセットといますか、安全性とセキュリティと一緒に語られることがあるのかなと思っています。ここでは安全性について特に触れられていないですが、AIを使うなら安全性ということも当然問題になるのかなということで、セーフティというのでもあっていいと思います。セーフティの一環としてセキュリティというのがあるのかなと私は思っているので、セキュリティだけを取り出して、セーフティが抜けているのは逆に何でなのだと。あえて落とすという選択肢もあるのかもしれませんが、セーフティというのを入れる対象になり得るかなと思っております。それが第1点です。

第2点が、5番目の公正競争確保の原則というところで、ブレットポイントの1番目に、特定の国や企業がデータの保有を集中させることにより云々（うんぬん）とありますが、この記載は国レベルの話と企業レベルの話がちよっと混じっているかなと思いました。

主権の侵害というのは、これは国家主権の侵害ということだと理解したのですが、特定の国がデータを独占すると、国家主権が侵害されるので、それが行われる社会であってはならない、と結構ハイレベルというかかなり政治的な問題の話がここでされているようにも思いました。

企業においては独占することによって、その支配的な地位を利用した不公正な競争を行ってはならないというのは、あり得る考え方かと思っていて、これから議論が始まる場所ですが、考えていかなければいけない問題であるということはそのとおりに思います。

もう一個は特定の国がデータを独占して、国家主権が侵害された、例えば日本の国家主権が侵害されるというのは、私は政府関係者ではないのでイメージが湧かないので、もうちょっと説明していただけたらなと思っております。【北野副議長】

企業に関しては、おっしゃるとおりですが、主権をなぜ入れたかという二、三十年ぐらい前は、アフリカまたは南米の地域がランドサットのデータを持っていないので、農業の収穫の状況というのが自分たちは把握できない。一方で巨大穀物メジャー等はランドサットのデータを持っていたので、そこで非常に安く買ったたかれるとか、情報のギャップによって基本的な国家産業に対してインパクトを与えるなど、それをどのような意味合いで主権というかどうかというのはありますが、かなり影響を与えるような情報ギャップによって、極めて不利を被ったという議論が、ユネスコの報告書にもあります。80年ぐらいに出ていると思います。

そこを主権という言葉を使うかどうかというのはありますが、同じようなことが、AIでも起きる可能性があって、国の基幹産業に対して自分たちで状況が把握できなくて、外の企業、またはほかの国が状況を完全に把握しているとなった場合には、現実的な主権行使が阻害され

るということもあり得るのではないかという意味で記載しております。

これは少し書きぶりも含めてご議論いただければと思っています。

**【須藤議長】**

どうもありがとうございます。

ただいま、北野副議長からおっしゃっていただいたとおりで、ここは重要な点ですが、書きぶりについては厳密な書き方等を考えなければならぬところもあるかと思えます。

例えば、イメージですが、エストニアに行ったときに、彼らはX-Roadで国民のトランザクション、政府のトランザクション、金融のトランザクション全部のデータがトランスペアレントになって扱いやすくなっているが、ブロックチェーンを使っていて、データ交換は水平的にやっている。

そのときに、我々が自信を持って言いたいのは、バックドアなんかを使って、政府が情報収集する国があるが、我々はそんなことは一切しないということを明確に、想定する国は言いませんが、そういう国は結構いろいろあるということで、それは頭に置いて書かなければいけないが、書き方は福岡構成員がおっしゃったように、十分留意しなければならないと思えます。

それから、セキュリティとセーフティの問題、私もなるほどそのとおりだと思いながら聞いていました。セーフティのことについてもきちんと触れた上で、やはり言うべきかと思えますので、考慮させていただきたいと思えます。

堀構成員。

**【堀構成員】**

主権の話は是非入れていただきたいと、やはり対中国にさらっとこういうメッセージが入っているのはいいことだと思います。

**【須藤議長】**

どうぞ、お願いいたします。

**【岩本構成員】**

エヌ・ティ・ティデータの岩本です。セキュリティのことで、実際のビジネスの中では、い

わゆる悪意に基づく犯罪者集団というのが大変大きな議論になっています。

国際的にもサイバーセキュリティに関する会議はいろいろされており、先ほど堀構成員もおっしゃったように、国際的協調の枠組みで撲滅という文言が入っているのは大変いいことだと思いますが、もっと強調してもいいと思います。書きぶりはもちろん注意するというのは先ほど議長がおっしゃったとおりです。中国でもいろいろな方々と話をしていますが、彼らは合法的に全部データを持っており、現地の中継センターを通すことによってアリペイにデータが流れていくという情報も取れております。誰が実際にマネーを送ったかなどの情報が即座に画面に出てくるぐらいのところまでできている状況です。

これはITとしては素晴らしいかもしれませんが、どう使うかということが重要であり、おそらくどのような世界でも悪意を持つ集団は必ず出てきてしまうため、いくら教育してもやめさせることは困難です。

我々の社会の中では、それに対して公的権力の行使が適切かどうかは議論を置くにしても、そういうことを防衛する仕掛けがもう少し必要だと思います。リアルの世界ではインターポールのようなものがありますが、サイバーの世界はもっと複雑になるため、我々と価値観を共有する人たちとの国際的競争の枠組みというものは必要だと思います。私もいろいろな所で言っておりますが、実現するとなると難しい問題も多い状況でもあります。

ワシントンではすでにサイバーウォーと言っており、少し次元が違うかもしれませんが、現実にはAIが動き出すと、AIがAIネットワークでつながるため、一つのAIだけで何かするのではない時代が必ず来てしまいます。どう犯罪から防御するかということは非常に社会の安定に寄与することになると思うため、もう少し踏み込んだ書き方ができるとよろしいかと思えます。

**【須藤議長】**

ありがとうございます。

ここら辺が、恐らく国際会議でも底流に流れる重要な論点だと思いますので、その点は十分考慮させていただきたいと思えます。ありがとうございます。

ほかありますでしょうか。

**【北野副議長】**

これに関連して、先ほどランドサット等の話をしましたが、もう一つのポイントはそのときにユネスコが出している報告書が、「Many Voices One World」というタイトルです。これは1980年にユネスコから出ています。”Towards a new more just and more efficient world information and communication order.”という内容です。

この報告書では、途上国の情報というのは、途上国はニュースエージェンシーを余り持っておらず、先進国のニュースエージェンシーは途上国の災害、犯罪、といったことばかり報道するので、途上国に対するイメージに非常にバイアスがかかり、それによって投資などが減退するといったことで、途上国が十分なボイスを持っていないという情報ギャップによって、著しく途上国が不利を被っているということが書かれています。

これは、正にいろいろな検索エンジンを使った場合であるとか、AIを使っていろいろなフィルタリングをしたときに起き得るバイアス、これがデータバイアスに基づくものと、それをフェイクニュース的なもので、インテンショナルにバイアスをかけていくということをマニピュレーションもできることになりますので、80年の段階ですでに素の情報のやり取りでこれだけ問題がでていて、それがAI的なことでアンプリファイされる危険性もあるということです。

今のところというのは、割と国家の枠組み的には非常に重要なポイントになってくるのではないかと思います。補足でした。

#### 【須藤議長】

非常に重要です。ありがとうございます。

そこは強調するべきだろうと思います。

ほかはいかがでしょうか。

#### 【武田構成員】

4章の全体構成ですが、(1)の人間中心の原則というのは、中身を読むとリテラシー、教育とか、社会の適正な利用とか、人間の創造性、イノベーションみたいなことが書いてあって、要は(2)から(7)までの話がここに凝縮して書いてあるという構成になっています。だったら人間中心というのを、大原則という格好で上に出したらどうかというのがまず一つです。

その上で人間中心とは何かということで、(2)の教育・リテラシーはもちろん大事なので第1の原則で、これは誰1人AIから取り残されないというようなメッセージにしてはどうかと思います。次の(3)～(6)のプライバシー、セキュリティ、公正競争、公平というのを見

ますと、これらは一塊にして第2の原則と言えないかと。プライバシーやセキュリティを考えないでどんどんやっちゃっている人が、プライバシーを一生懸命考えている人に比べて得をするという、それは公正の一つの問題ですし、(3)～(6)までを一つの塊にして競争が公正に行われる社会、というようなことにしてはどうかと思います。それから、(7)のイノベーションというのは、これは斬新で非常にいいと思いますので、人間中心の第3の原則として、人間の創造力が発揮され続ける社会のようなメッセージにする。

A Iの日本が考える大原則は人間が中心なんだというのが最初のメッセージに来て、その人間中心とは何かと言うと、教育・リテラシー、これはSDGsみたいな言葉で言うと、誰一人取り残されない、A Iから取り残されない社会なんだというのが第1の原則。第2の原則が、ともするとA Iというのはずるして得する人がいかにないわけですから、公正な競争がされる社会というのが第2点。第3の原則が人間の創造性というものが発揮され続ける社会、そんな格好の大括りで、まず3原則のようなことで展開すると、よりメッセージ性が出るのではないかと思います。以上です。

#### 【須藤議長】

おっしゃるとおりかもしれません。これは一つの考え方でして、4のA I原則というタイトルですが、人間中心のA I社会原則論みたいな、言い方は検討しなければいけないと思いますが、何かそうしてもいいのかもしれませんが。(1)はまえがきと言うか、大命題として上げておくという手もあるかもしれません。

こちら辺について、構成案、何かご意見があれば。

平野構成員、お願いします。

#### 【平野構成員】

今の議長のご提案に賛成でございます、この会議のタイトルをもう一度、社会原則の会議なので、机上資料2の社会原則と開発原則が平行ですが、むしろ社会原則はもっと大きく、それを支えるパーツの一つが開発原則のようにした方が、この会議が正にそこに集中して議論をしてきましたし、もしかしたらいいのかなと思います。

その中で、先ほどのご発言の人間中心とか、正に会議のタイトルが人間中心でございますから、それは相当大きなところに来るという構成がいいのかなとちょっと思います。

【須藤議長】

ありがとうございます。

北野副議長。

【北野副議長】

そういうたてつけというのは十分可能だと思います。

【須藤議長】

ありがとうございます。

近藤構成員、お願いします。

【近藤構成員】

近藤です。

持続性のある社会というところに入るのかどうか。戦争のない平和な社会というのを目指してみんな頑張っていると思うので、何かそういうメッセージが、この中のどこかに入るのかなと思いました。

【須藤議長】

これも国の名前を出しますが、中国が今年の夏にAIの大原則を出したときに、軍事利用というのは明確に書いています。フランス政府はこの3月に軍事イノベーション庁を設立しましたが、AIの利活用、軍事的利活用を最も重視すると書いています。

これは近藤構成員のおっしゃるように、我々が望ましいと思っていない方向に持っていくものなのですが、これはリアルな世界の状況も考えないといけない一方で、抑止力とか理念としてそれをうたうかうたわないかというのは、極めて重要な論点ですが、ご意見はいかがでしょうか。皆様のご意見を伺っておきたいと思います。

【安宅構成員】

今、守りしか書いてないですよ、セキュリティ云々（うんぬん）のところは。攻めのところは何をやろうと攻める人は攻めるので、基本的には裏で行われるということを考えると、ここに書くことはある種我々の手足を縛ることにしかならず、むしろ別の外交的、あるいは物理

的パワーによる抑止の方が大事だと思います。

**【須藤議長】**

ありがとうございます。

私も、今置かれた状況からすると、触れるとかえって、今、安宅構成員がおっしゃったように我々が縛られてしまって、言うことを聞いてくれない巨大国家が勝手に動いて、どうしようもなくなるということもあり得るので、ここら辺は慎重に構えないといけないと思います。

近藤構成員、何かご意見は。

**【近藤構成員】**

男の人たちの発想だなとすごく思います。AIという新しい技術が人だけではできないことに、さまざまな新しい可能性の領域を開いてくれるということにもし期待できるのであれば、世界中の人が一生懸命目指すべきは、そちらなのではないかと。

でも、現実に悪い人たちがたくさんいたり、悪いことを考える国もあったりする、でもその悪いという歴史的なこれまでの世界の歴史を考えれば、何かもう少しそこに新しい可能性があるといいなと思います。

**【須藤議長】**

ありがとうございます。

重要なお意見だと思います。これも国際的な議論が必要なところで、ここはOECDやその他の国際会議に出れば、書き込むか書き込まないか極めて重要な判断を迫られることになると思います。頭には置いておかなければいけないと思いますが、記載するかどうかについてはちょっとペンディングさせておいてください。

どうぞ。

**【中川構成員】**

今の点について、IEEEのEADでもかなりAIウェポンのことを書き込んでいるのですが、その論調は基本的には人間がコントロールできない状況のものをつくってはいけないということだけが、コンセンサスとしてあるようで、要するに人間が責任をとれなくなってしまうのは、偶発的な衝突とかいろいろ危ないと。絶対的に人間がコントロールできる形で行けとい

うことだけは世界的なコンセンサスがある。そこから先はむしろ政治の問題だと考えて、I E E EのE A Dなんかは書いているようです。

**【須藤議長】**

ありがとうございます。

正に、中川構成員がおっしゃるようにながりの政治的に高度な判断を伴いますので、このところはペンディングでお願いします。考えなければいけない重要なテーマであることは間違いありませんが、慎重な取扱いは必要だと思います。

どうぞ。

**【安宅構成員】**

これは難しすぎる問題なのですが、先ほど岩本構成員からあったように、撲滅云々（うんぬん）のところ、地球防衛軍的なサイバーセキュリティ軍みたいなものをつくるという提案はいいかもしれない。攻めを止めるみたいなものは単純に我々の自由度を破壊する可能性があるもので、そっち側（がわ）の方が効力がある気がします。

男の子的な発想なのかもしれませんが、ゲームを見れば分かるとおりに、少なくとも我々のテストステロンか分からないが、何かの本能はそっちの方向にながり衝動があることは間違いありません。ただ、事実なので、受け入れた中で、現実の解を探るのがいいと思います。

**【須藤議長】**

ありがとうございます。

私の立場からすれば、恐らく北野副議長の立場もそうですが、国際会議に訴えて、合意形成を目指さなければいけない重要なものです。これはG 20にかけなければいけないということになります。

そのとき、合意ができないアジェンダを入れることは極めてリスクがあるので、この点をご了解いただきたいと思います。

**【北野副議長】**

A Iの軍事利用について一般的に非常に問題だという議論がされているというのは承知しています。それと同時に、A Iの国際会議などでは、A Iを使った武器の方が精度は高いので、

市民に対するダメージは減らせるという議論もあって、これに反論するのは極めて難しい議論でもあるわけです。

ですから、そういうところも冷静に見た上で、どちらかと言うと、ここは軍事、ナショナルセキュリティ以外のところを出した方が、合意できると思います。軍事、ナショナルセキュリティの議論は、軍縮関係など、別の議論の場があるので、こちらで触れないでも良いだろうと思います。

【須藤議長】

ありがとうございます。

丸山構成員。

【丸山構成員】

軍事の話からちょっと外れますが、先ほど北野副議長のMany Voicesの話を聞いていて、いわゆるテクノロジーの民主化ということがもう少し書き込まれてもいいかなというような気がします。包摂というのが理念の中にあるとすれば、この技術は誰もがアクセスできるようにするという原則がどこかにあってもいいような気がします。

【須藤議長】

ありがとうございます。

私も大賛成です。テクノロジーの民主化という概念は今の議論の上位概念に来ると思いますので、かなりここは重視して書くべきだろうと思います。

昨日、マイクロソフトのCEOが日本で講演したようですが、AIの民主化はマイクロソフトの理念になっていて、いろいろ戦略的な意味があるのですが、ここは字面というか表の論理では、かなり合意形成がかなりできそうなので、押していけるころだと思います。

浦川構成員。

【浦川構成員】

AIを活用する上で、ここ4、5年当社でもかなりAIを実装しておりますが、データという点についてももう少し取り上げるべきかどうか、皆様のご意見をお聞きしたいです。しかる

べきデータを食わせればある程度正しい方向を示してくれるのがAIだということを、実装していて実感しております、実際に幾つかのシステムでAIを実装しておりますと、AIを正しく動作させるために、データのノーマライズをすとか、選別をすとか、足りないデータを保管すとか、自社だけで足りないデータをどこから持ってくる、エコ化みたいなものですが、そういったことに関するワークロードの方が圧倒的に実はかかっておりまして、これはAI-Readyのキーワードなのか、若しくはAI社会原則のところなのかもしれませんが、やはりデータの重要さというのが、これまでの情報処理システムで必要であったデータと、AIを活用して、人間を豊かにするために必要なデータというのは、実はよく見ると、随分性質やドメインが違う、というところが実際にAIの実務に携わっている身からすると実感しておりますので、この辺りに触れるかどうかはお任せしたいと思いますが、気になったところです。

【須藤議長】データの扱いについて極めて重要で、これもマスコミだけでなく、国際会議でも議論になっているところで、何らかの形で触れられればと思います。また、ユヴァル・ノア・ハラリ氏著『ホモ・デウス』においても今後の人類はデータ中心主義で、人間存在の今の定義がみんな破壊されると書いてありますが、そういう議論も頭に置いて、書くべきことかなと思います。これは人間の在り方にも関わる場所なので、きちんと何らかの形で書き込めればと思います。

#### 【浦川構成員】

もう1点、先ほどの議論で、AIの利用は人が自ら判断し、という議論があったかと思いますが、保険屋の立場ではないですが、実際に実装してみて、今、何が起きているかと言うと、当初はAIが判断したものを目検で最終チェックをしてお客様に契約をオーダーするというのをやっていたのですが、簡単な契約については、自動判断の方が正確度が上がって、人間の判断を上回ってしまったので、比較的簡単なケースは保険の契約にしても、支払いについても、アルゴリズムの目検チェックはするのですが、一件一件のトランザクションはもう機械に任せるという領域が出始めています。ですので、そこに対して、1件1件の人の判断なのか、それを司るベースとなるアルゴリズムを含めたシステム実装全体に対して人間が責任を負うのかといったところは、技術の発展によって少しずつ変わっていくものなのかなと実装して感じております。ちょっと補足しました。

【須藤議長】

ありがとうございます。

そういうリアルな認識も必要です。銀行などでも、与信の審査もかなり人工知能を使って信頼するというケースが増えているということはメガバンクの役員の方から伺っていますので、かなりそれが一般的になりつつあると思います。

もう時間がございません。全体を通して、本当に本日も重要な、生産的なご意見を多く頂きましたが、何かあればおっしゃっていただきたいと思います。

【新田参事官】

事務局からよろしいでしょうか。

先ほどのデータの関係の議論ですが、岩本構成員からのご議論で、これについては今のAI-Readyの社会の中のイノベーションシステムのところで、あらゆるデータが安全かつ新鮮、そういったところで利用できるというところにしっかり書き込んでいくと、ある程度いけるのかなと考えております。

それから、このAI原則に関してですが、今回、事務局の方でも骨子を構成する上で、先ほど来、議論が出ていますが、これから価値観を共有するほかの国の方々とどう合意形成をしていくのかということを考えますと、この原則はできるだけシャープに表現するという事で考えております。

そういう意味では、人間中心のところも確かに、まえがきに少し全体的な大原則みたいなことを書いているので、ちょっと重複感があるのかもしれませんが、人間中心はブレットが二つありますが、やはり人間を代替し能力を拡充し、お互いに高度化できるみたいな、そういうものでなければならない。より具体的なメッセージで各国と調整したいと考えています。2番以下も同様でございます。

それから、AI開発利用原則に関して、冒頭に少しご議論がありましたが、これもEUにおきまして、今はまだ、AI for people、これは新聞記事の情報でしかないので、中身がどのようなものが込められていくのか現時点では分かりませんし、ディスカッションの中にもございましたように、新聞記事が正しいのであれば、損害責任保険の加入を義務づけるなど、保険の議論なんかもなされているということを、今、このAI社会原則、基本的には名宛人が社会、特に政府や自治体とそういったところを対象としています。

それで、今後のEU、冒頭赤石からも紹介がありましたが、EUのハイレベルとも少しディ

スカッションを始めているという状況の中で、EUがどういうことを考えているのか、政府や社会の原則か、民間の関係者に対してはどう考えているのかというところ、もう少し情報を収集して、我々のAI原則をどう構築していくのかというところは、もう少し情報を収集しながら、この4.2の扱いをどの程度にしておくのがいいのかは、もう少し時間をかけさせていただければと考えております。情報をできるだけ速やかに入手いたしまして、少し議論を含めさせていただければと考えているところでございます。

事務局からは以上です。

**【須藤議長】**

ありがとうございます。

今、最後の方におっしゃったことで言うと、12月上旬ぐらいまでの状況を把握させてほしいということですね。書き方はかなりそこで、国際会議で合意をとるための条件づくりが必要になるので、状況認識が必要になるということです。したがって、皆様をお願いする執筆もかなりそのあたりの影響は受けるということをお願いいたします。

それから、今の新田参事官のご発言は4のところの大見出し、人間中心のAI社会原則とは書かない方がいいということの意味しているのですか。

**【新田参事官】**

人間中心は原則中の一番上ということで、いの一番に言ってもいいかなという意味ですが、ふわっとした原則ではなくて、英語で言うとオーギュメンテーションやコーディベロップメントとか、そういうものではなければならないと、はっきりした原則で構成するのがよいのではないかと、そういう趣旨です。

**【須藤議長】**

人間の能力の拡張、あるいは人間の進化と言っていいと思うが、そういうものをまず最初に打ち出すということですね。

ほかは何か。

どうぞ。

**【安宅構成員】**

最初の3ページ目にあった基本理念のサステナビリティの話をもうちよっと入れてもいいのではないかと考えていまして、サステナブルな未来に向けた利活用という視点が欲しいというのと、それに向けた危険な知見を共有するという話があつていいのではないかなど。

というのは、最近、『銃・病原菌・鉄』で著名なジャレッド・ダイヤモンド教授の『Collapse（邦題：文明崩壊）』という本を読んでいて、マヤ文明にしてもイースター島にしても、オーストラリアにしても、人間がよかれと思ってきたことは大体人間を滅ぼしてきたという過去の歴史が繰り返されていて、何も悪意なく人間は滅んできたということ、これらを防ぐためにこのAIがあるんだということを打ち出すことは結構いい話ではないかと考えています。

**【須藤議長】**

それは、このサステナビリティについてもうちよっと重要な情報を4で書き込んではいかがでしょうかということですか。

**【安宅構成員】**

人間中心の原則にサステナビリティを入れてしまうというのは一つあると思います。これが自己中ではない人間中心である。未来の世代、3世紀後の未来まで含めた人間中心。

**【須藤議長】**

サイモンという経営学者が言っているように、人間、限界合理性というのがあって、見えないというのが人間の特徴なわけです。そこを突破するツールとしてもAIは道具としては使える可能性は極めて強い、可能性があるということは書き込めるかなと思います。

**【羽鳥構成員】**

その他でお話ししようと思いましたが、先ほどのお話で、AIのいろいろな画像診断がこれから非常に進んでくる。病理の判断が、AIがこれは良性である、病理のすぐれた先生はこれは悪性であると、余りすぐれていない病理の先生は、例えば遠隔医療で、AIの診断を信じて良性だと。しかし、3年後に実はガンであったとか。同じように、くも膜下出血の原因になる、未破裂脳動脈瘤でも同じようなことが起こると思います。

AIでMRAの診断をしたが、そうではないとレベルの低い先生が判断してしまった。そのときに先ほどの話だと、医師が最終的な責任をとるという考えと、浦川構成員がおっしゃって

いたように、製造者責任を持たせるという考えもあって、今の議論だと保険で賄うという話もあったかと思いますが、大分現実的な話になってきて、どう説明していったのかと悩むところでは。

【須藤議長】

ここら辺ももう少し、ブレインストーミング、メールベースでもやらなければいけない可能性もあり得るし、もう一回皆様のご都合を確保して議論しなければいけない可能性もありますが、つくった人、利用する人の共同責任という概念もあるわけです。それと保険適用の仕方とかあるし、中川構成員がおっしゃるように、もう一つ、二つ、それもおっしゃっているわけですが、AIの能力を認めて、未来発展可能性を認めると、人格的な要素を議論しなければいけないという可能性もあるし、そこら辺はまだ今意見は並行しているので、議論する必要があるかもしれないです。

北野副議長。

【北野副議長】

今の議論で言うと、AIの人格といったときに、例えば保険の話で、保険でカバーできるタイプのパニッシュメントというか、特に民事的なところだったり、刑事的なところをやらなければいけなくなると保険でカバーできるという話でもなくなってくるので、これをどうするかというのは非常に大きな問題です。パニッシュメントが起こるということで、納得する人々という、そういう状況がある場合。

もう一つ、先ほどの4.2の立ち位置なのですが、今は社会原則とパラレルに置いているのですが、先ほど平野構成員の方からあったように、社会原則の下に開発利用原則とする場合には、社会原則のところから紐づける書き方をする必要があって、その場合には、インターガバメントの枠組み、または政府の施策として開発利用原則というものをエンカレッジするという、そういう紐づけをどこかで書くということになると思います。そういう方向の書きぶりでもいいのか、それともこれは自主的にするので、余り紐づけたくないとすると、社会原則とパラレルに置いた方が紐づけなくて済みます。どちらにするかというところが一つの悩みどころかと思えます。

【須藤議長】

これも事前のミーティングで北野副議長、事務局、私で議論したところで、ストーリーのつくり方が難しいです。ここはもう一回議論した方がいいかもしれないですね。

【新居構成員】

はじめにとおわりのところで、先ほどもご指摘があった国際的な枠組みにどう参加していくかみたいところで、言及があると思います。ムーブメント、コンセンサスという言葉を使っていて、もちろんムーブメント、コンセンサスを起こしていこうというのは大事なのですが、その先に期待している日本という国の立ち位置の期待感というのがあるはずで、せっかく出すのであれば、そこまで踏み込んでしまうことで、ムーブメントを起こした先に日本がこういう立ち位置になっていくんだということまで、この原則で宣言できると、ただ頑張るぞみたいな終わりにならないというところが結構大事かなという感じがしました。

【須藤議長】

正に、これは国際会議だけではなくて、Society 5.0を実現するということを高らかに言っていて、それをもっと具体的にここで考えているよということを行いますので、それはおっしゃったように強調するべきだろうと思います。

平野構成員、お願いします。

【平野構成員】

責任の賠償責任、民事責任、刑事責任の話で、私の直感では余り細かくここでは決めない方がいいかなと思います。まだ時間が足りないというか、世界的にもまだ詰まっていない。

例えば、民事の保険の話も保険でカバーすればいいというだけではなくて、不法行為の専門家として言わせていただくと、要するに、非難すべきではない人にとりあえず責任を課して賠償で保険というのは、私は納得できません。

例えば、自動車、悪くないのに原因が分からなくて、自動車のメーカーの責任にされちゃうというのは、僕はすごく不満です、不法行為法としては。ただ、被害者を救済するという社会保障としてはよく分かる。そこは切り分けなければいけないとか、実はすごく論点もあって、まだそこは解が出てないので、このビックピクチャーを我々が描いている中では、今後検討する必要があるぐらいの表現の方がいいのかなと。先ほどの、ワトソンと名指ししてはいけな

いですが、ワトソンと~~医者~~医師のどちらの責任というのも、ワトソンは「医師が最後は判断してください」と言っているわけで、それは当然裏に賠償責任があるわけですね。ここは、ここで決めることができないぐらい、今後議論というか、検討すべきぐらいの方がいいのかなというのが私の直感でございます。

**【須藤議長】**

ありがとうございます。

非常に重要なご意見です。書きぶりはかなり気をつけたいと思います。おっしゃるとおりだと思います。

どうぞ。

**【中川構成員】**

今の平野構成員のご意見に少しだけ追加させていただくと、そういった責任といろいろ考えるときに、AIの説明能力、イクスプレイナビリティというのは非常に絡んでくるので、その二つが絡んでいるということが分かるように書いていただくといいのではないかと思います。

**【須藤議長】**

ありがとうございます。

そのとおりだと思います。そこはうまく、記述いたしましょう。皆様のご協力が必要です。

もう時間が来ました。ご予約もあると思いますので、今日の議論はこれにてやめたいと思いますが、非常に皆さん、コンセプトが明確になってきて、いい議論をしていただきました。これをベースにして、事務局で取りまとめて、皆さんにまた今後の執筆等でご相談することになります。

事務局、今後のことについてお願いいたします。

**【新田参事官】**

資料2をご覧ください。今のところ、第8回で原則案を取りまとめ、公表というスケジュールを守りたいと思っておりますので、事務局の中の関係もありまして、1カ月ほどスケジュールが押していて申し訳ございませんが、12月の次回の会合で公表を目標と考えています。

執筆チームの皆様にも、ご協力を是非ともよろしくお願ひしたいと思ひます。

なお、来年には、ユネスコハイレベル会合、G20、G7とございますので、しっかりこのA I原則案を世界に発信できるようにと考へてございます。

次回の会合につきましては、執筆チームの作業状況を見ながら別途事務局からご案内させていただきますことになります。どうぞよろしくお願ひいたします。

また、本日の資料、薄いですが、郵送をご希望される方は机上に残していただければ結構でございます。

事務局からは以上です。

**【須藤議長】**

どうもありがとうございます。

いまありましたように、何とか12月の完成を目指したいというのが事務局の立場です。これもよく分かります。1月以降重要な国際会議がたくさんありますので、その前にやはりまとめて打って出なければならぬという政府のお立場もございますので、この点、かなりプレッシャーがかかると思ひますが、短時間で、執筆のご協力をお願ひ申し上げたいと思ひます。

今の取りまとめ案から大幅に分量が増えるということはないですよ。

**【新田参事官】**

分量は、ある程度作業を進めながらで、別段何ページなければいけないとか、そういった制限はございません。

**【須藤議長】**

また、北野副議長、須藤、事務局で今日のご議論を踏まえまして、執筆者案をまとめますので、ご協力のほどお願ひ申し上げます。

それでは、これにて本日は閉会といたします。

どうもありがとうございました。

-閉会-