別添６

報告書（案）の
別紙２

整理した論点に関する意見募集に寄せられた主な意見[1]
に対する考え方

番号	意見の要旨	意見に対する考え方
ＡＩの定義について		
１	ガイドラインの対象範囲を明確にし、ソフトウェア開発等の萎縮を防ぐために、ＡＩやＡＩシステムの定義を明示すべきである。	「ＡＩ」の定義については、ＡＩ研究者の間でも多種多様な定義が示されており、確立した定義が存在しないことなどから、意見募集に当たっては「ＡＩ」の定義を示すことは控えてはどうかとしていたが、開発ガイドラインの対象範囲を明確にするために、開発ガイドラインの目的に照らして必要な限りにおいて、開発ガイドラインにおける用語としての「ＡＩ」の定義を示すこととした。 具体的には、開発ガイドラインにおける用語としての「ＡＩ」を「ＡＩソフト」と「ＡＩシステム」とを総称する概念とした上で、前者を「データ・情報・知識の学習等により、利活用の過程を通じて自らの出力やプログラムを変化させる機能を有するソフトウェア」と定義する一方で、後者を「ＡＩソフトを構成要素として含むシステム」と定義し、その例として、ＡＩソフトを実装したロボットやクラウドシステムを挙げている。
ガイドラインの対象範囲となる主体について		
２	ＡＩ（ソフト）単体では、制御可能性、セキュリティ、安全性等を確保することは困難であり、ＡＩ（ソフト）開発者に過大な負担を課す恐れがあることから、ＡＩシステム又はＡＩネットワークの開発者が留意すべき事項を中心にガイドラインを定めるべきである。	開発原則においては、ＡＩ ソフトのレベルではなく、ＡＩシステムのレベルにおいて透明性や制御可能性等を確保するため、ＡＩシステムの開発者が留意することが期待される事項を定めることとした。
３	開発ガイドラインの対象範囲となる人的主体についてより詳細に定義すべき。具体的には、研究者と開発者を分離すべき。また、開発者と提供者を分離すべき。	今日のＡＩの研究開発においては、研究と開発が一体的・連続的に行われることが多くなっており、研究者と開発者を区別することは必ずしも容易ではない。 なお、本ガイドラインの対象とする開発の範囲は、学問の自由の尊重、社会に与える影響の大きさ等に鑑み、閉鎖された空間（実験室、セキュリティが十分に確保されたサンドボックス等）内での開発は対象とせず、ネットワークに接続して行う段階に限定している。
４	ＡＩネットワークシステムを開発・設計する主体は、利用者ではな	ＡＩシステムの研究開発（ＡＩシステムを利用しながら行う研究開発を含む。）を行う者（自らの開発したＡＩシステムを用いてＡＩネットワークサービスを提

[1] 提出された意見の全部について、次に掲げるＵＲＬのウェブサイトに所掲。
＜http://www.soumu.go.jp/main_content/000490102.pdf＞

	く、開発者に当たるものとして整理すべき。	供するプロバイダを含む。）を「開発者」と定義した。 なお、他人が開発したＡＩシステムを構成要素とするＡＩシステムを開発する者（当該ＡＩシステムによりＡＩネットワークサービスを提供するプロバイダを含む。）は、当該他人との関係においては「利用者」と位置付けることが適当である。
ガイドラインの対象範囲となるＡＩ及び開発の段階について		
５	自律型ＡＩと非自律型ＡＩを区別し、既に普及が進んでいる後者のみをガイドラインの対象範囲とすべきである。	開発ガイドラインは国際的に参照される非拘束的なソフトローとして策定されるべきものであり、ＡＩに関する技術が予想を超えて加速度的に発展していることや、欧米の政府や民間等におけるＡＩの開発原則等の検討においても「非自律型ＡＩ」のみならず、今後急速な発展が見込まれる「自律型ＡＩ」も念頭に置いたもの有力となっていることに鑑み、「自律型ＡＩ」の発展をも見据えて開発ガイドラインを策定することが適当であると思われる。
６	ガイドラインの対象範囲をＡＩの機能に即して限定すべきである。	限定的な機能を有するＡＩシステム（特化型ＡＩ等）であっても、ネットワークを通じて他のＡＩシステムと連携することなどにより、高度で多様な機能を実現し、広範な利用者及び第三者に便益及びリスクを及ぼす可能性がある。したがって、その機能如何にかかわらず、ネットワーク化され得るＡＩシステム、すなわち、ネットワークに接続可能なＡＩシステムを本ガイドラインの対象範囲としている。
７	ガイドラインの対象範囲となる「開発の段階」から除外される研究開発を画定する「閉鎖された空間」の定義を明確にし、クラウドコンピューティング時代に即した研究開発体制を考慮すべきである。 具体的には、仮想閉域網などで論理的に閉鎖された空間をこれらの定義に追加すべきである。	クラウド等情報通信ネットワークシステムを利用したＡＩの研究開発に配慮して、開発ガイドラインの対象範囲から除外される「閉鎖された空間」における研究開発の例に、「セキュリティが十分に確保されたサンドボックス等」を加えた。
開発原則の構成及び順序について		
８	人間の尊厳と個人の自律の尊重を内容とする倫理の原則は、他の原則の基礎となる根底的な価値を定めるものであり、最上位の原則として位置づけるべき。	開発原則は、あくまでも開発者が留意することが期待される事項を機能的に整理したものであり、原則の順序は、価値の序列を示すものではない。 なお、開発ガイドラインの依拠すべき価値を列挙する基本理念において、第一の基本理念として、「人間がネットワーク化されたＡＩと共生することにより、その便益がすべての人によってあまねく享受され、人間の尊厳と個人の自律が尊重される人間中心の社会を実現すること」を掲げることにより、開発ガイドラインが人間の尊厳と個人の自律が尊重される人間中心の社会を実現するという理念に立脚したものであることを明確にした。

透明性の原則について		
９	技術や分野の特性を踏まえて、透明性の基準を検討すべきである。	技術的中立性に鑑み、透明性の原則の解説において、「採用する技術の特性や用途に照らし合理的な範囲で、ＡＩシステムの入出力の検証可能性及び判断結果の説明可能性に留意することが望ましい」との説明を加えた。
１０	機械学習における深層学習 のメカニズムは、少なくとも研究者にとってはブラックボックスではなく、研究者が何を入力して、どう学習させたのかが理解可能であれば、入出力や判断の透明性は確保されている、という解釈をして良いのか、という点を明示すべき。	開発者にとってＡＩシステムの入出力の検証及び判断結果の説明が可能であり、かつ、開発者が利用者等に適時適切にアカウンタビリティを果たす用意ができている場合には、ＡＩシステムの透明性（入出力の検証可能性及び判断結果の説明可能性）が留意されているものと理解するのが適当である
制御可能性の原則について		
１１	ＡＩを安全に停止できるかどうかは明確ではなく、制御可能性の原則は実現性が低い。制御可能性の限界を踏まえ、受入れ可能なリスクとして社会的に受容できるようにするための原則を策定すべきではないか。	制御可能性の原則を含め開発原則は、開発者が遵守すべき一定の基準を示すものではなく、開発者が、留意して対応し、対応状況についてアカウンタビリティを果たすことが期待される指針を示すものである。したがって、開発者は、制御可能性の原則に鑑み、採用する技術の特性に照らして可能な範囲でＡＩシステムの制御可能性に留意するとともに、制御可能性の原則への対応の在り方について、制御可能性の限界やその程度も含め、アカウンタビリティを果たすことが期待されると理解するのが適当である。
国際的な整合性の確保について		
１２	セキュリティや安全性等に関しては、国際的に参照されている工業標準等と調和するガイドラインを作成する必要がある。	関連する原則の解説において、国際的に参照されている標準や規格等を参照することが開発者に期待される旨を記述した。
１３	ネットワークに起因するリスクに対応するため、SCM 関連規格等を参照して、原則を見直すべきである。	関連する原則の説明文等において、国際的に参照されている標準や規格等を参照することが開発者に期待される旨を記述した。
１４	各原則に対し実現を求められる範囲・水準の国際的な整合を図るべき。	各原則に対し実現を期待される範囲・水準の国際的な整合を図る観点から、関係するステークホルダに期待される役割として、各国政府、国際機関、開発者、市民社会を含む利用者など多様なステークホルダ間の対話の促進、ベストプラクティスの共有等を掲げた。

関係するステークホルダに期待される役割等について		
１５	開発原則の実効性確保に関して、開発者が自発的に提供する情報に基づいて第三者機関が開発原則への適合性を評価し認証する制度の創設が例として挙げられているが、このような制度は、開発者の自由な開発を萎縮させ、ＡＩの発展を妨げる可能性を有するものであり、ガイドラインに書き込むことには極めて慎重であるべき。	あたかも開発者に認証を受けることを義務付けるなど公的な認証制度を創設するかのような誤解を招くことのないよう、意見募集時の資料において開発原則の実効性を確保する方法の例として掲げていた「開発者がその開発するＡＩに関し自発的に提供する情報に基づき、公正中立で高度な専門性を有する第三者機関が当該ＡＩの開発原則への適合性を評価して認証する制度」に関する項目を削除した。
１６	各原則について、その実現例・満足例（ホワイトリスト）を作成し、公開すべきである。また、ホワイトリストの作成に当たっては、実証実験を可能とする特区等の指定、マルチステークホルダーによる議論の場を設定し、実用例に基づいたホワイトリストの作成・改良を行うとともに、社会合意を進めるべきである。	ＡＩに関する技術や利活用の分野に応じて開発原則への適切な対応の在り方が異なり得ること、またＡＩに関する技術が加速度的に発展していくことが見込まれることに鑑みると、（分野共通）開発ガイドライン自体に「ホワイトリスト」を盛り込むことは適当ではないものと思われる。 一方、開発原則への適切な対応の在り方については、各々の開発者が利用者等にアカウンタビリティを果たす上で自主的に判断すべきことを基本としつつ、関係するステークホルダに期待される役割として、「標準化団体等は、本ガイドラインに適う推奨モデルを作成し公表することが期待される」と記載した。
１７	開発者のみによるリスクの抑制には限界があり、利活用の場面において利用者が留意し対応すべき事項も念頭に置いてガイドラインを作成すべきである。	開発者と利用者との責任の適切な分担に留意しつつ、開発ガイドラインの検討を進めるとともに、利用者が留意して対応することが期待される事項についても検討を進めることとしたい。
１８	利活用ガイドラインは、事業者のみを名宛人とし、一般消費者は対象とすべきではない。また、利用者の信頼・期待の保護については、市場原理に委ねるべき。	利活用ガイドラインについては、市場における公正な競争を通じた利用者の利益の確保を基本に据えて議論を行い、利用者が留意し対応することが期待される事項について、利用者の種別に配慮しつつ検討していきたい。