とりまとめ骨子（案）

令和 8 年 11 月 18 日

人工知能戦略専門調査会

制度の在り方検討ワーキンググループ

I. ＡＩ法成立以後の情勢

1. リスクの情勢

○　ＡＩ法制定時から、重大リスクへの将来的な法規律の必要性が示唆されていたところ、令和８年以降、AI を悪用したサイバー攻撃、化学物質、生物、放射性物質 、 核 、爆発物 （以下「CBRNE」という。）等、ＡＩの予期せぬ挙動による重大な危害の発生への懸念が増大している。

2. 諸外国における制度整備等

○　各国でＡＩの安全性への懸念が高まりつつあるものの、法制度に基づく規律による対応が先行しており、米国カリフォルニア州法や EU 法において、AI を開発・提供する事業者に対して一定の義務が課せられている。

3. 我が国における政府等の取り組み

○　広島ＡＩプロセス等、国際的な協力の枠組みを推進しつつ、ＡＩ法に基づく「適正性確保に関する指針」や各省庁におけるガイドライン等を整備し、ＡＩ法においても国による調査研究を規定されているが（第 16 条）、同法に基づき調査を行ったものの、一部の事業者からは調査への協力が得られなかった事例も存在する。

○　ＡＩＳＩ（ＡＩセーフティ・インスティテュート）においては、ＡＩモデルの安全性評価等を実施している。

II. 制度の在り方

1. 今後の AI に関するリスクへの対応の全体像

(1) 基本的な方向性

○　ＡＩが経済社会の発展の基盤となる技術であることは令和７年のＡＩ法制定時から変わりなく、ＡＩの開発や利活用そのものを規制することは妥当ではない。引き続き、ＡＩの研究開発及び活用の推進を図っていき、我が国ＡＩ産業の国際競争力を向上させていくという国の基本方針を変える必要はない。

○　ＡＩに関する様々なリスクについても、原則としては、可能な範囲で既存の個別法やガイドラインに基づいて対応していくという、現行のＡＩ法の対応枠組みを維持しつつ、一方で、現行法令で対応できないような未知のリスクやこれまでにない規模で危害が発生するリスクが具体化している分野については、限定的に、ＡＩを活用した情報処理サービスを提供する者（以下、「ＡＩサービス提供者」という。）に対して、法令による規律を設け、ＡＩによってもたらされる重大な危害の発生を防止し、対処していくことが妥当である。

○　ＡＩによってもたらされる重大な危害が発生し又は発生するおそれがある場合には、政府が事業者から必要な情報を迅速かつ的確に収集し、これにより重大な危害の拡大や発生の防止のために政府がとるべき措置を検討することが不可欠である。

○　必要な情報を収集するための国の権限や事業者の義務及びこれを担保するための罰則に関する規定を設けることが必要であると考えられるが、あくまで、ＡＩの研究開発や利用

に係る作為や不作為に対する直接的な罰則ではなく、行政からの報告の求めや措置の勧告に従わなかった場合に罰則を科す、間接的な罰則規定を設けることが妥当である。

○　一方で、グローバルに事業を展開し諸外国の法令を遵守しなければならないＡＩサービス提供者の負担も考慮し、諸外国の規律との調和・整合も考慮しつつ、我が国における法整備を進めるべきである。

○　高性能なＡＩモデルの開発に係る技術は、膨大な投資を要する競争領域であり、事業者にとっても、企業秘密として社外に開示することがためらわれる情報も多数ある。したがって、事業者の保護を要する情報にも配慮した運用を行うことの必要性も十分に検討されるべきである。

○　また、ＡＩサービス提供者の負担も考慮し、法的な義務を課す場合には予見可能性、基準等の明確性等を高めるとともに、届出や報告等の事務手続きについても、他法令に基づく手続きとの重複を避けるなど、不要な事務負担を生じさせないよう配慮すべきである。

(2)　対象となるリスクの整理

○　ＡＩに関する多様なリスク全てにＡＩに関する横断的な法令で対応することは妥当ではない。個別の領域について生じる問題は、引き続き、原則として個別領域における既存の法令やガイドライン等により対応を図っていくべきである。

○　一方で、本ワーキンググループが主に検討の対象としてきた、令和７年のＡＩ法制定以降に特に顕在化した、ＡＩがサイバー攻撃、ＣＢＲＮＥテロ等の分野で悪用され又は予期しない挙動を引き起こし、社会に重大な混乱を引き起こすリスクの類型については、ＡＩにより、社会全体への悪影響が著しく増幅され又は創出され、社会に重大な悪影響を及ぼす事態が惹起されることから、ＡＩサービスの安全性確保に着目した法令により対応することが妥当であると考えられる。

(3)　対象となる事業者の整理

規律の対象となる事業者については、以下の二つの観点から特定することが考えられる。

①　ＡＩの性能の観点

上記のようなリスクの整理を踏まえると、規律の対象とする事業者は、原則として、社会に重大な悪影響を及ぼす事態を惹起する上で中核的な役割を果たし得る能力を有するＡＩを提供する事業者に限定するのが妥当である。

なお、特定の用途に特化しているなどの理由から、その仕様上、上記のリスクを生じる可能性が小さいと考えられるＡＩサービスについては、規律の対象から除外することを検討すべきである。

②　国内における利用の観点

重大な悪影響を及ぼし得るＡＩサービスについて我が国の法令においてなんらかの規律を課すとしても、あくまで国内法である以上、対象とするＡＩサービスはその重大な悪影響が国際社会全般に生じるものではなく、我が国で発生する可能性があるものに限定する

必要がある。また、我が国の国内法令による規律の対象となっており法的義務を負っていることを事業者が明確に認識し、法令が事業者に対し安定的に適用される観点からも、国内での利用状況を考慮することが合理的である。他方、国内において一定程度利用されていれば、海外事業者のＡＩサービスについても、当然に規律の対象とすべきである。

また、他社が開発するＡＩサービスを自社のサービスに組み込んで活用する場合などについて、サイバー攻撃やＣＢＲＮＥテロ等の（2）で念頭に置く特定のリスクの危険性を防止するためには、上流のＡＩサービス提供者において安全確保措置を講じることが合理的である。他方で、既に安全確保措置の実施が法律で担保されている他社が開発するＡＩサービスを自社のサービスに組み込んで利用している企業等について、一律に上流の事業者と同様の義務を課すことは、イノベーションの促進の観点からも慎重に検討すべきである。

2. ＡＩの安全性確保に着目した法令の具体的な内容

上記を踏まえ、法律において、以下の事項を定めることが妥当である。

- ○ ＡＩによるリスクから安全を確保するための措置を講じる義務
- ○ 国に情報を報告する義務
  - ・ 重大な危害の発生に関する報告
  - ・ 報告徴収
- ○ ＡＩによるリスクから安全を確保するための措置に関する国の勧告・命令権限
- ○ 報告義務や命令に違反した場合の罰則
- ○ ＡＩＳＩ(ＡＩセーフティ・インスティテュート)の役割の拡充