国際的な規範やその基礎として我が国が先んじて策定した「人間中心のAI社会原則」を踏まえ、AIの研究開発・活用における**適正性確保の考え方、適正性確保のための基本方針**を提示。その上で、**主体別に特に配慮すべき事項**を整理した形で構成。

# 【指針骨子(たたき台) 全体構成案】

- 1 適正性確保に関する基本的な考え方
- ●適正性確保の考え方

人間中心	プライバシー保護	セキュリティ確保	公正競争確保
公平性、安全性	透明性、アカウンタビリティ	リテラシー	イノベーション

●適正性確保のための基本方針

リスクベースでのアプローチ	アジャイルな対応
ステークホルダーの積極的な関与	一気通貫でのAIガバナンスの構築

- 2 研究開発機関及び活用事業者が特に配慮すべき事項
- ∨透明性の確保と誠実なアカウンタビリティ ∨徹底した安全性確保 ∨持続可能なイノベーションの実現
  - 3 国民が特に配慮すべき事項
- ✓能動的なリテラシーの習得・応用
  - 4 国及び地方公共団体が特に配慮すべき事項

eeイノベーションの強力な推進 ee社会全体における A  ${f I}$  リテラシーの向上 ee行政としてのアカウンタビリティ ${f 1}$ 

## (参考) AIリスクへのアジャイルな対応(イメージ)

- 事業者・国民の自主的かつ能動的な取組を促すAI法第13条に基づく指針 (AI指針) を整備。
- 今後、AI指針やAI事業者ガイドライン等を基にした、AIセーフティ・インスティテュート(AISI)による AIモデルの技術的評価の実施、当該評価も踏まえたAIがもたらすリスクに係るAI法第16条に基づく 調査研究、各府省庁ガイドライン等を有機的に組合せて、アジャイルに対応。

## New

#### AI法第13条 指針

評価・調査を踏まえた アジャイルな見直し

AIの適正性確保のための自主的かつ 能動的な取組を促す基本的な考え方



## 各府省庁ガイドライン等

分野ごとの具体的な対策等

ガイドライン等を基とした 評価・調査

To-Be

AISIによるAIモデル技術的評価

AIの安全性にとどまらない適正性に係る評価

(As-Is) 評価観点に基づく評価ツールの提供



## AI法第16条 調査研究等

国民の権利利益の侵害事案について実態把握

適正性の評価・調査

自主的取組の促進



## 事業者・国民による自主的取組

政府の取組に積極的に協力し、自分事として適正性確保に取り組む

#### 政府による支援

リスク対策の技術開発等

