AI規制論の選択肢

- 焦点リスク：生成AIの前後いずれを念頭に置くのか
  - 情報処理AI：製品安全とプロファイリング（EU AI法当初提案）
  - **情報生成AI**：偽・誤情報と情報環境全般への影響（EU AI法汎用目的AI条項）
- 対象：後者なら、基盤モデル全般と、**巨大な基盤モデル**のいずれを念頭に置くのか
  - ＋偽・誤情報が流通するプラットフォームの責務の在り方
- 手法：巨大な基盤モデルなら、どの規制強度を念頭に置くのか
  - 法に基づかない要請のみ（自主規制）
  - **技術情報やリスク軽減策の透明性義務（自主規制＋透明性）**
  - **リスク評価・軽減・体制整備義務（共同規制）**
  - 具体的な行為規制やモデル承認義務（直接規制）
- 制度運用：自主・共同規制規律の具体化手法はいかなるものか
  - 整合規格、行動規範、 AI Safety Instituteの役割

EU AI法の汎用目的AIモデル条項

- 汎用目的AIモデル（生成AIはその典型）提供者の義務（53条）
  - 設計や学習等の技術文書作成と当局への提供
  - 下流事業者への情報開示
  - DSM指令4条（学習データ権利者オプトアウト）等のEU著作権法遵守措置
  - 学習データの詳細な要約の公表（欧州委AIオフィスがテンプレートを作成）

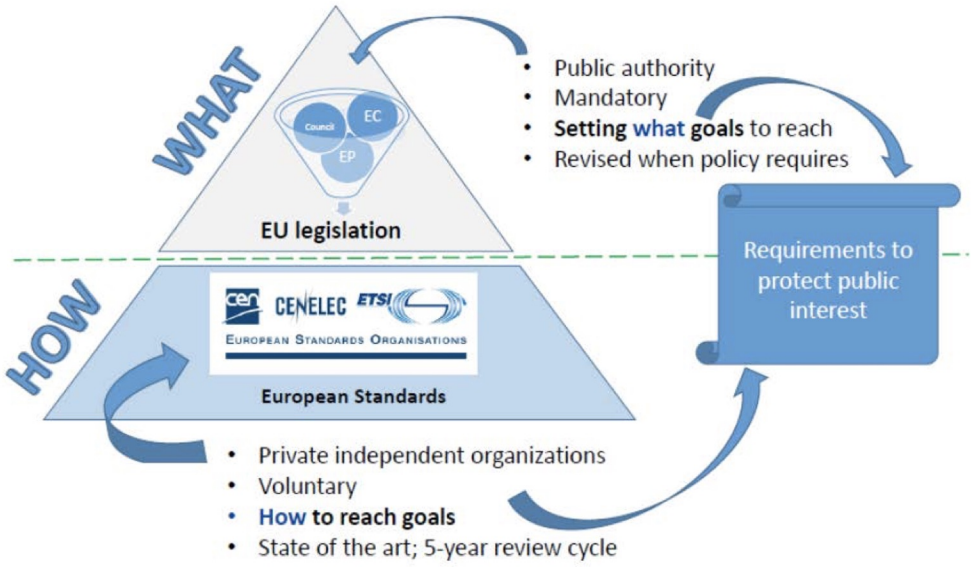
  ※AI生成コンテンツに対する機械可読マークの付与（50条2項）
- システミックリスクを伴う汎用目的AIモデル（10^25FLOPS等）提供者の義務（55条）
  - レッドチームテスト実施を含むモデル評価
  - **システミックリスクの評価・軽減**
  - 重大インシデントへの対応文書化と当局への報告
  - サイバーセキュリティ対策

→整合規格（harmonized standard）と行動規範（codes of practice）による具体化

3条(65)「「システミックリスク」とは、汎用目的AIモデルの高インパクト能力に特有のリスクであって、**その影響範囲の広さにより連合市場に重大な影響を及ぼし、または公衆衛生、安全、治安、基本権もしくは社会全体に対する実際の若しくは合理的に予見可能な悪影響**により、バリューチェーン全体にわたって大規模に伝播し得るリスクをいう」

EU AI法と整合規格・行動規範

55条2項：システミックリスクを伴う汎用目的AIモデルの提供者は、整合規格が公表されるまでは、本条第1項に定める義務の遵守を証明するために、56条の意味における行動規範に依拠することができる。欧州整合規格への準拠は、当該規格がこれらの義務をカバーしている限りにおいて、提供者に適合の推定を与える。(53条4項にも同趣旨条文)



CEN-CENELEC "Drafting Harmonized Standards in support of the Artificial Intelligence Act (AIA)" (2022.3) より

56条（行動規範）
1. AI オフィスは、国際的なアプローチを考慮しつつ、本規則の適切な適用に貢献するために、EU レベルでの行動規範の策定を奨励し、促進するものとする。(...)
3. AIオフィスは、汎用目的AIモデルの全ての提供者と関連する国内の管轄当局に、行動規範の策定に参加するよう要請することができる。市民社会組織、産業界、学界、下流提供者や独立した専門家などのその他の関連する利害関係者は、このプロセスを支援することができる。(…)
6. AI オフィス及び欧州人工知能会議は、参加者による行動規範の目的の達成及び本規則の適切な適用への貢献を定期的に監視及び評価するものとする。(…)欧州委員会は、実施法により、行動規範を承認し、域内で一般的に有効とすることができる。（…）
7. AI オフィスは、汎用目的AIモデルの全ての提供者に、行動規範を遵守するよう要請することができる。(…)