ブリュッセル効果への対応
日本企業はEU-AI法にどう備えるべきか

2024年12月11日(水) 12:00-13:00

【プログラム】

12:00-12:05	開会挨拶：村上明子（日本AIセーフティ・インスティテュート）
12:05-12:25	EU-AI法の概要と日本企業が留意すべき対応の要点：古川直裕（株式会社ABEJA）、吉永京子（慶應義塾大学大学院）
12:25-13:00	パネルディスカッションとQ&A パネリスト： 実積寿也（中央大学） 工藤郁子（大阪大学） 古川直裕（株式会社ABEJA） 村上明子（日本AIセーフティ・インスティテュート） 吉永京子（慶應義塾大学大学院） 司会：東京大学（江間有沙）

UTokyo | TOKYO COLLEGE

東京大学未来ビジョン研究センター
Institute for Future Initiatives
The University of Tokyo

AI Center UTokyo

ブリュッセル効果（？）への対応：日本企業はEU-AI法にどう備えるべきか

EU-AI法の概要と日本企業が留意すべき対応の要点
前半部分

吉永　京子
慶應義塾大学大学院 政策・メディア研究科 特任准教授
東京大学未来ビジョン研究センター 客員研究員
ジョージタウン大学法科大学院テクノロジー法政策研究所ノンレジデント・フェロー

This research was supported by the JST Moonshot R&D project, JPMJMS2215

＜関連著書＞

- 商事法務「別冊NBL EU AI法概説」今年度刊行予定。
- 商事法務NBL 1269(2024.7.1)号～1278(2024.11.15)号まで9回に渡り「EU AI法」連載。
- 古川直裕・吉永京子「責任あるAIとルール」
  （一般社団法人 金融財政事情研究会、2024年5月） *Kindleも。
  責任あるAI（生成AI含む）とそれを実現するためのルール、人・企業・国がAIとどう向き合っていくべきかを考える手がかりとなるような一般の方向け書籍。
- 「EUのAI法と新興技術規制への視点」【特集：科学技術と社会的課題】、三田評論、2024年8月5日
  https://www.mita-hyoron.keio.ac.jp/features/2024/08-4_2.html
- 福岡真之介・杉浦健二・古川直裕・木村菜生子編著
  『AIプロファイリングの法律問題-AI時代の個人情報・プライバシー』（商事法務、 2023年10月）

＜参考＞

- 中崎尚「生成AI法務・ガバナンス―未来を形作る規範」（商事法務、2024年7月）
- 「EU AI規則の概要」欧州連合日本政府代表部、2024年９月
  https://www.eu.emb-japan.go.jp/files/100741144.pdf

1. EUのAI法概要

2021年4月　欧州委員会AI規則案を公表。
2023年12月　欧州委員会、EU理事会、欧州議会による非公式交渉（トリローグ）
2024年5月21日　成立
2024年7月12日　官報に掲載
2024年8月1日　発効

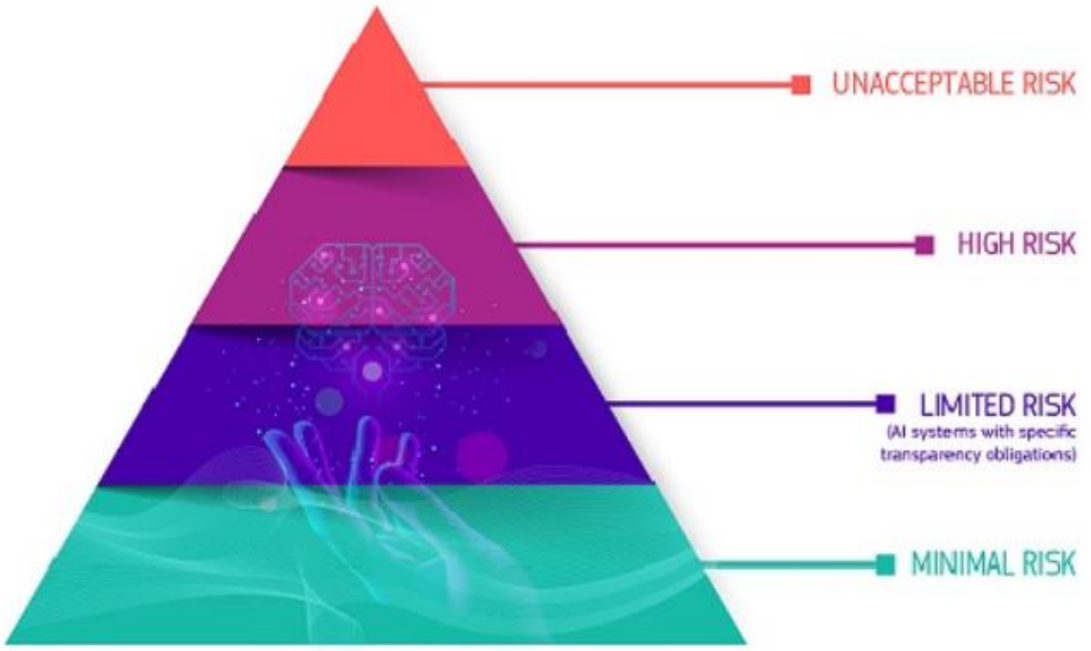
●包括的には**ハードローアプローチ**

- 世界で初めて**包括的にAIを規制**する法律が成立。Regulations（規則）は直接適用。

➢New Legislative Framework（NLF)の一環
➢27の加盟国の市場の統一
　⇒市場内で流通
➢EUの価値
➢法的確実性

- リスクベースアプローチ

「リスク」とは、害が発生する確率とその害の重大性の組み合わせを意味する。（EU　AI法第３条2号）



https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

2. EUのAI法の規制対象

規制対象	例
（1）禁止されたAI行為	サブリミナルな技法や操作的な技法、年齢・障害等の脆弱性情報を用いて人間の行動・意思決定に悪影響を及ぼすこと、ソーシャルスコアリングを行って人間に害を及ぼすこと、プロファイリング等のみに基づいた犯行予測、インターネット等からスクレイピングした顔認証DBを作成すること、生体情報分類システムの使用、公共の場でのリアルタイムの遠隔生体識別システムを法執行目的で使用すること（例外あり）
（2）ハイリスクAIシステム	リスクマネジメントシステムの構築や関係者への情報開示等の透明性の要求事項あり。適合性評価を受けること。
（3）特定のAIシステム	プロバイダーとデプロイヤーに、（ハイリスクAIよりは軽微な）透明性の要求事項 例：エンドユーザが接しているのはAI（チャットボットやディープフェイク）であることを通知すること
（4）汎用目的AIモデル (general-purpose AI model; GPAI)	技術文書の作成・更新、汎用目的AIモデルをAIシステムに統合するプロバイダー向けの情報・文書の作成・更新・提供、EU著作権指令遵守のためのポリシーの実行、汎用目的AIモデルの事前学習に使用されたコンテンツの概要の作成及び公開、域内代理人の指名。 システミックリスクがあるGPAIプロバイダーは上記＋モデル評価、EUレベルでのシステミックリスクの評価及び軽減、重大なインシデントの追跡・報告、サイバーセキュリティ保護の確保等の義務。（市場投入前のR＆D目的のモデルは対象外）

プロバイダー：開発者・提供事業者、デプロイヤー：利活用事業者・ユーザー事業者

3．適用範囲（第2条）

適用対象　（2条1項）

(a) EU域内で、AIシステムを市場投入、サービス提供している、または汎用目的AIモデルを市場投入しているプロバイダー（EUに所在しているかは問わない。）

(b)　EU域内に所在するAIシステムのデプロイヤー

(c)　アウトプットがEU域内で利用される場合、第三国に所在するAIシステムのプロバイダー及びデプロイヤー

(d)　AIシステムの輸入業者、流通者

(e)　自らの製品とともにまたは自らの名義や商標のもとにAIシステムを市場投入、サービス提供している製品製造者

(f)　EU内に設立されていないプロバイダーの認定代理人。

(g)　EU域内に所在し影響を受ける者（person）

部分適用（2条2項）

6条1項に従ってハイリスクAIシステムと分類されたもの（付属書Iセクション B に列挙）は、6条1項（ハイリスクAIの定義）、102条～109条（関連法令のハネ改正）、112条（見直し規定）が適用される。57条 （AI規制サンドボックス）は、AI法に基づくハイリスクAIシステムに対する要件が、当該EU調和法に統合されている場合にのみ適用される。

適用対象外（2条3項以下）

・軍事、防衛、国家安全保障
・他国の公的機関、国際機関による利用
・科学的な研究開発（R＆D）
・市場投入・サービス開始前のAIシステムまたはAIモデルの研究、テスト、開発行為（実環境でのテストは除く）
・私的利用 (purely personal non-professional activity)
・フリー・オープンソース・ライセンスでリリースされたAIシステム

中崎尚「生成AI法務・ガバナンス－未来を形作る規範」（商事法務、2024年7月）、古川直裕「EU AI法概説（第1回）総則と禁止AI（1条～5条）」NBL1269(2024.7.1)号、「EU AI規則の概要」欧州連合日本政府代表部、2024年９月

4. 罰則（99条～101条）

違反は多額の罰金

例えば、

違反行為	罰則
5条（禁止されたAI行為）違反	最大で3,500万ユーロ、または違反者が企業である場合は前会計年度の全世界年間売上高の7%までの行政罰のいずれか大きいほうが適用（99条3項）
5条ではない事業者や、認定適合性評価機関に関連するプロバイダー、認定代理人、輸入者、販売者、デプロイヤーの義務や特定AIのプロバイダーおよびデプロイヤーの透明性義務規定に違反	最大で1,500万ユーロ、または企業である場合は、前会計年度の全世界年間売上高の3%までの行政罰のいずれか大きいほうが適用（99条4項）

5. 適用時期

113条

施行時期	項目
2026年8月2日	下記以外
2025年2月2日	一般規定 (Chapter 1)と禁止されたAI行為（Chapter 2)
2025年8月2日	認定適合性評価機関 (Chapter 3 Section 4)、汎用目的AIモデル (Chapter 5)、ガバナンス (Chapter 7)、78条（秘密保持）、101条以外の罰則 (Chapter 12)
2027年8月2日	6条1項と対応する義務

2025年5月2日、
Code of Practice

https://artificialintelligenceact.eu/implementation-timeline/

Implementation Timeline

Timeline of key dates

Last updated: 1 August 2024

Items in blue relate to the Application of the Act.

Download timeline as image

2024

Date		Related AI Act Content
12 July 2024	The AI Act is published in the Official Journal of the European Union. This serves as the formal notification of the new law.	Article 113
1 August 2024	**Application:** Date of entry into force of the AI Act. At this stage, none of the Act's requirements apply—they will begin to apply gradually over time.	Article 113
2 November 2024	**Member States:** Deadline for Member States to identify and publicly list the authorities / bodies responsible for fundamental rights protection, and to notify the Commission and other Member States.	Article 77(2)

YOU ARE HERE

2025

Date		Related AI Act Content
2 February 2025	**Application:** Prohibitions on certain AI systems start to apply (Chapter 1 and Chapter 2).	Article 113(a) Recital 179
2 May 2025	**Commission:** Codes of practice shall be ready by this date.	Article 56(9) Recital 179
2 August 2025	**Application:** The following rules start to apply: • Notified bodies (Chapter III, Section 4), • GPAI models (Chapter V), • Governance (Chapter VII), • Confidentiality (Article 78) • Penalties (Articles 99 and 100)	Article 113(b)

6. 適用までの間「AI協定」（AI Pact）

Overall framework to prepare the AI Act's implementation

Pillar I	Pillar II
Engagement with all stakeholders through **webinars**	Voluntary pledges for **companies** of any size
Open to all types of organisations	*Focused on providers and deployers*
Organisation of **webinars** open to all stakeholders to address specific aspects of the AI Act. Public, private, NGOs, etc	Set of **voluntary pledges** calling on organisations to proactively work towards taking steps towards the implementing some of the key provisions of the AI Act
Exchanges with stakeholders on how the AI Office can effectively help organisations in the implementation of the AI Act	Workshops specifically tailored to the **companies**

https://digital-strategy.ec.europa.eu/en/policies/ai-pact

- 2024年9月25日の欧州委員会のプレスリリースによると、100社超の企業が自主的誓約に署名と発表。

欧州委員会のプレスリリース：
https://ec.europa.eu/commission/presscorner/detail/en/ip_24_4864

- ITのほか、電気通信、医療、銀行、自動車など幅広い産業のEU企業が参加。
- 参加している企業の例：Accenture, Adobe, Amazon (Amazon Europe Core), Google, IBM, Mastercard, Microsoft, NEC, OpenAI, Salesforce, Vodafone, Samsungなど。

JETROのプレスリリース：
https://www.jetro.go.jp/biznews/2024/09/4ccc193a8488579a.html
AI Pact, European Commission: https://digital-strategy.ec.europa.eu/en/policies/ai-pact

7. 禁止されたAI行為（5条）

施行時期: 2025年2月2日

禁止されたAI行為	備考
(a) サブリミナルな技法や操作的な技法を用いたAIシステム	サブリミナルとは潜在意識にこっそり働きかけること
(b) 脆弱性情報を用いて人間の行動や意思決定に悪影響を及ぼすこと	脆弱性＝年齢、障がい、社会経済的状況
(c) ソーシャルスコアリングをおこなって人間に害や不当な扱いを与えたりすること	ソーシャルスコアリング＝社会的行動や個人的な特性に基づいて個人や集団を評価分類すること
(d) プロファイリング等のみに基づいた犯行予測	犯罪行為に直接関係する客観的で証明可能な事実に基づいて行う犯罪行為に関する人間によるアセスメント支援AIには適用されない。
(e) インターネット等からかき集めた（スクレイピングした）顔認証データベースを作成すること	
(f) 職場や教育の現場における感情認識	医療や安全上の理由を除く
(g) センシティブな個人情報を収集する生体情報分類システムの使用	人種、政治的意見、労働組合への加入、宗教または思想、性生活および性的嗜好を推測・推論するために生体データに基づき個人を分類するもの
(h) 法執行目的での公共の場でのリアルタイム遠隔生体識別	例外：不明者の捜索、生命・身体的安全に対する特定の重大かつ緊急の危機・テロ攻撃に対する危機防止、付属書２に定める最大4年以上の身体拘束刑に該当する犯罪等に対する犯罪捜査の実施、訴追、刑罰執行のための被疑者の所在の特定等。

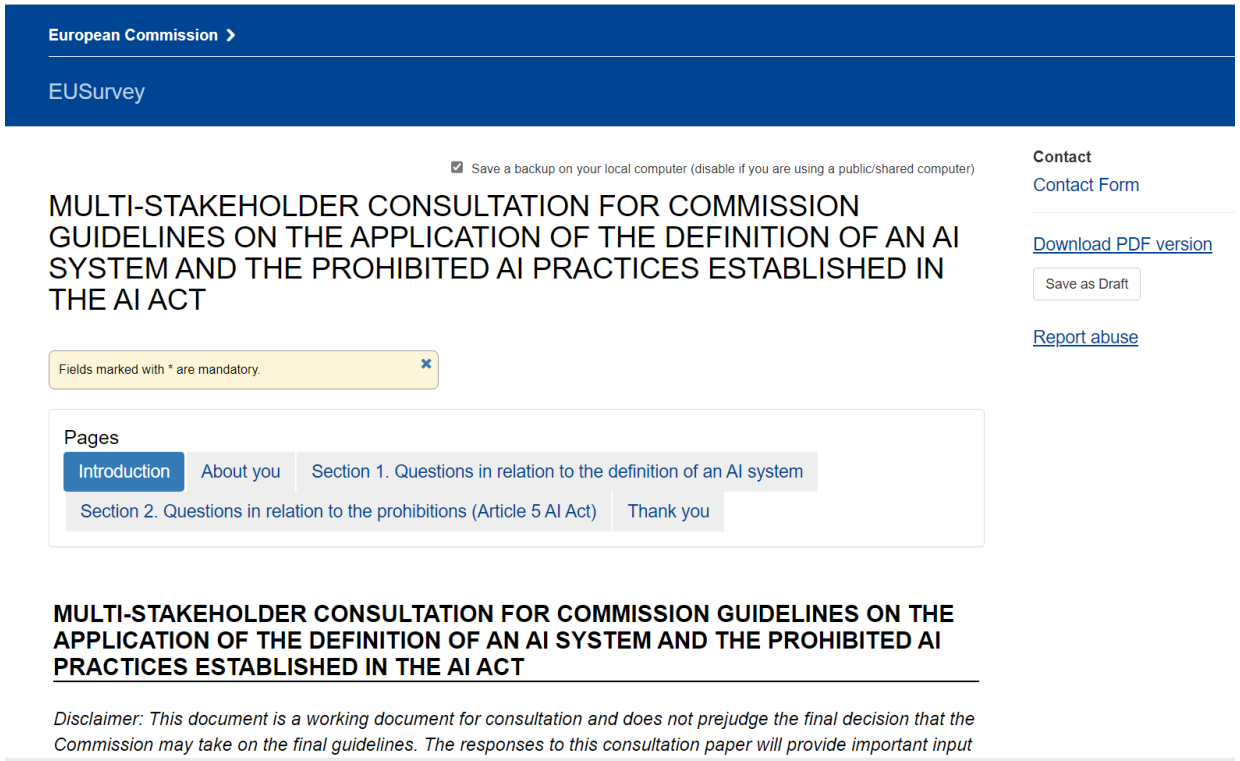
7. 禁止されたAI行為（5条）

施行時期: 2025年2月2日

EUがAIシステムの定義と禁止されたAI行為に関する意見提出を受付中。
期限：11 December 2024 (till 23:59)

MULTI-STAKEHOLDER CONSULTATION FOR COMMISSION GUIDELINES ON THE APPLICATION OF THE DEFINITION OF AN AI SYSTEM AND THE PROHIBITED AI PRACTICES ESTABLISHED IN THE AI ACT

https://ec.europa.eu/eusurvey/runner/Prohibitions-and-Definition-Survey-2024



8. ハイリスクAIシステム（6条～49条）

施行時期: 2026年8月2日
（6条1項は2027年8月2日）

- **ハイリスクAIとは？**

  人間の健康・安全や基本的権利に重大なリスクをもたらしかねないAI

  ・第６条１項の(a)(b)号に定める条件を満たすAIシステム（附属書Ⅰ）

  ーAIシステム自体が製品または製品の安全要素である(a号）、法によって第三者認証の対象となるもの(b号)

  ・付属書IIIで言及されているAIシステム（第６条第２項）8分野：1)生体識別、2)重要インフラ、3)教育および職業訓練、4)雇用・動労者の管理と自営業へのアクセス、5)必要不可欠な民間サービスおよび公的サービスへのアクセスと享受、6)法執行、7)移民・亡命および国境管理、8)司法の運営および民主的プロセスに関するもの

- **ハイリスクAIのプロバイダーに求められていること**（第８条～第15条）

リスク管理システム（９条）、データガバナンス（10条）、技術文書（11条）、記録保持（12条）、透明性確保とデプロイヤーへの情報提供（13条）、人間による監督（14条）、AIの正確性・堅牢性・サイバーセキュリティの確保（15条）

- **プロバイダー・デプロイヤー等の義務**（16条～27条）

プロバイダー：品質管理システム（17条）、文書の維持、ログの保存（19条）、システム要件不適合やリスク対応（20条）等

デプロイヤー：第26条（技術的・組織的措置の実施、人間の監督、適切な入力データ、リスク対応、重大インシデント報告、ログの保存等）

- **支援制度**

  －AI規制サンドボックスに参加することも可能。

ブリュッセル効果への対応：日本企業はEU-AI法にどう備えるべきか

EU-AI法の概要と日本企業が留意すべき対応の要点
後半部分

株式会社ABEJA　弁護士　古川直裕

1．特定AI

- 注意点→　ハイリスクAIにも適用がある！
  実はピラミッドではない・・・
- 類型1　人間と直接的に会話するAI
  AIと会話していることが分かるようにする　ただし、常識的にわかる場合は別
- 類型2　コンテンツ生成AI
  機械可読な形で人工的に生成されたと検出できるようにする
- 類型3　感情認識または生体分類AI
  関連EU法に従ってデータを扱い、その旨の通知
- 類型4　ディープフェイクAI
  人工的に生成されたことを開示
  ただし、芸術、フィクション等とわかる場合は作品を害さないようにDFの存在を明示
- 類型4´　公共の利益に関する事項を公共に情報提供するためのテキスト生成AI
  人工的に生成されたことを明示
  ただし、人間が最終編集していれば例外

2．General Purpose AI

- 通常GPAIとシステミックリスク付GPAIの分類
  ①適切な技術ツールや手法での評価
  　学習の合計計算量が10の25乗FLOPSだと推定がかかる
  ②政府による指定
- 通常GPAIの義務　無料のオープンソースライセンスには適用なし（SR付は別）
- 政府への情報提供義務（技術文書）
  学習・検証・テスト用データに関する情報、アーキテクチャとパラメータ数、モデルと学習過程の設計仕様、学習時間や学習に用いた計算資源、推定消費電力など・・・
- 下流への情報提供義務
- 著作権遵守ポリシーの策定
- 学習に使ったコンテンツに関する十分詳細な概要（テンプレが政府から提供される）
- COPを守れば上記の義務の遵守が推定される

３．GPAI with systemic risk

- システミックリスク付GPAIの義務
  普通GPAIの義務に加えて・・・
  標準化されたプロトコルによるモデル評価
  システミックリスクのアセスメントとリスク軽減
  政府への適切な報告
  適宜のサイバーセキュリティー保護
- 　COP遵守による遵法性推定

４．Code of Practice

- COPについての定めもあり
  重要な点としては・・・
  国際的なアプローチを考慮してCOPを定める・・・！！！
  COPが具体的な目的をさだめ、目的達成の手段（適当な場合はKPIも）を示す・・・！！！

５．取引先による遵法要求

- 取引先が適用されることで影響を受けることも
- ハイリスクAIの開発受託
  10条のデータ義務
  12条で要求されているログ取得
  15条の精度など・・・
- GPAIの開発受託（基盤モデルでなくても）
  情報開示に必要な情報の提供、ドキュメント作成など・・・
- 事前に作業として工数と費用を見積もっておく
- AI法の理解

６．今後も続く一致体制

- COPと同じような「従えば法律を守ったことになる標準」が今後も作られていく。
  ①リスクマネジメントシステム
  ②データガバナンス
  ③記録取得
  ④透明性と情報提供
  ⑤人間の関与
  ⑥精度
  ⑦頑健性
  ⑧サイバーセキュリティ
  ⑨品質マネジメントシステム
  ⑩適合性アセスメント

大阪大学

ブリュッセル効果への対応：日本企業はEU-AI法にどう備えるべきか

EU AI法「行動規範 (Code of Practice)」について

大阪大学 社会技術共創研究センター 特任准教授
工藤郁子

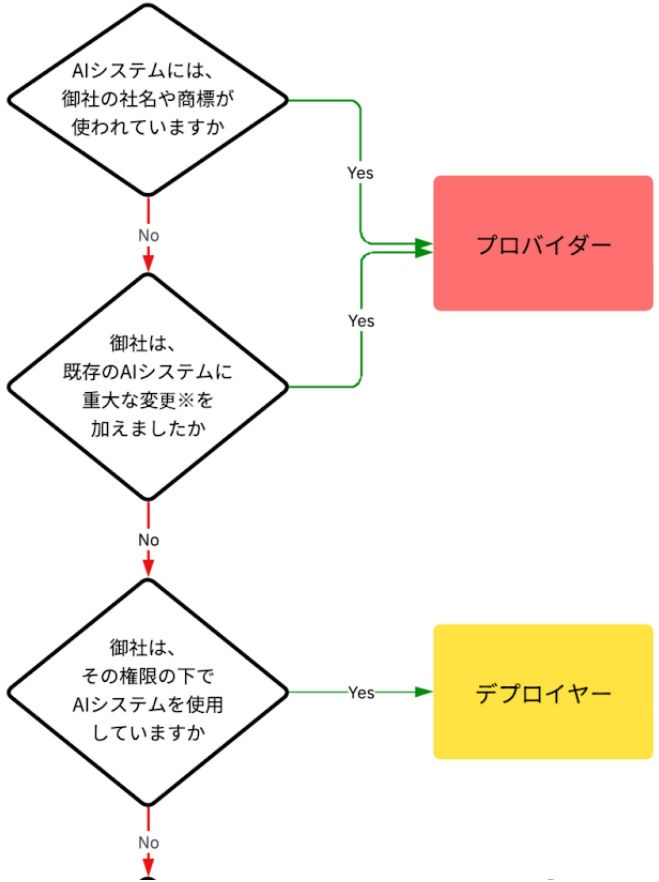
2024年12月11日

大阪大学 社会技術共創研究センター
Research Center on Ethical, Legal and Social Issues

Osaka University
Research Center on
Ethical, Legal and
Social Issues

EU AI法と行動規範

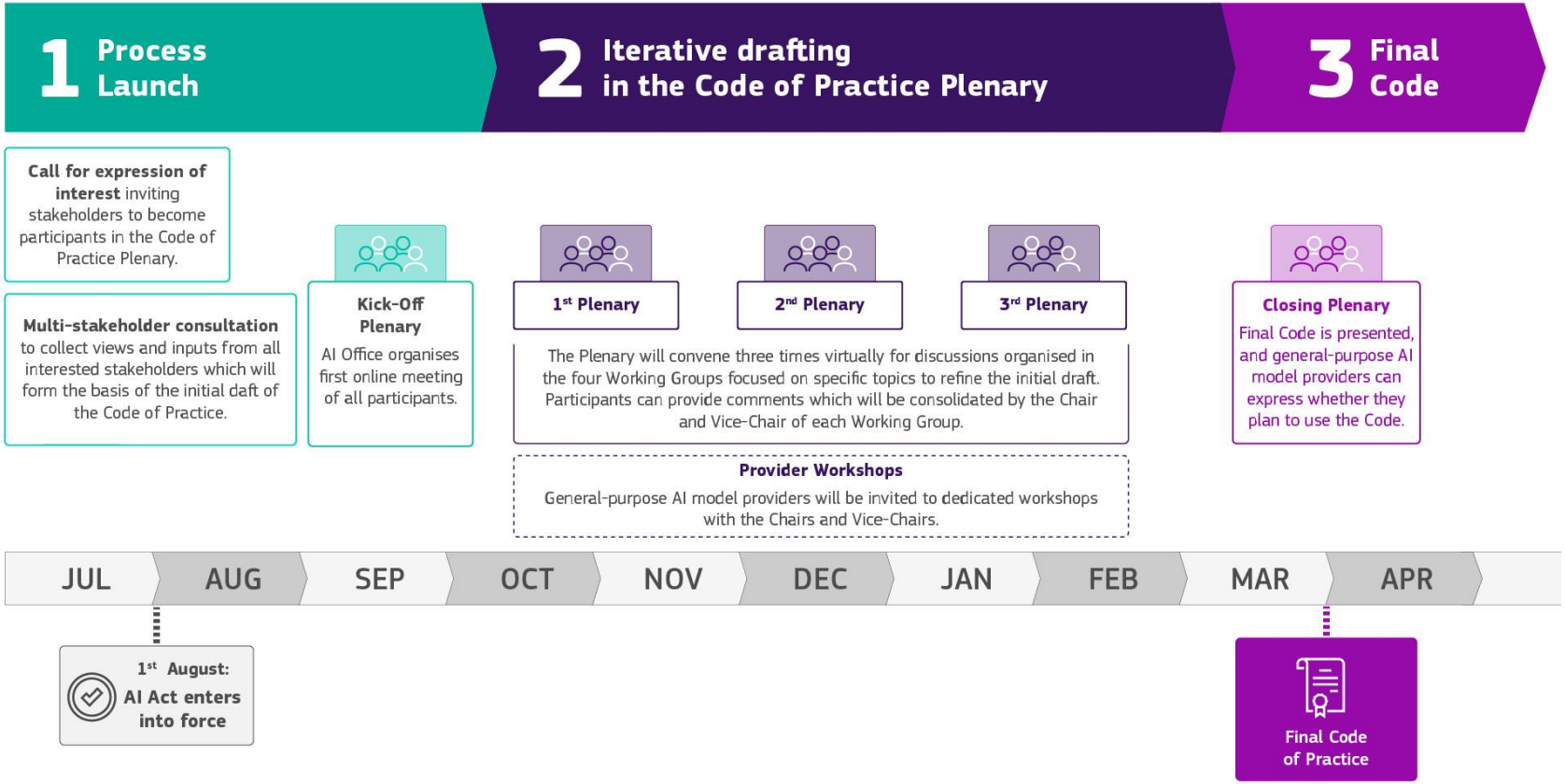
- EU AI法では「行動規範（CoP: Codes of Practice）」と呼ばれるガイドラインを策定することになっている（56条）
- 汎用目的AI （GPAI: general-purpose AI）モデルに関する義務について法解釈を具体化するもので法的拘束力はないが、行動規範への準拠は「法令遵守の推定」として機能する
- 汎用目的AIモデルのプロバイダーだけでなく、汎用目的AIモデルを利用する下流プロバイダーにも影響し（53条・55条参照）、域外適用も
- 2024年8月1日のAI法発効後、行動規範は、マルチステークホルダー協議を経て、2025年4月までに策定されるスケジュール



（via https://citadel-ai.com/ja/blog/2024/02/16/eu-ai-what-it-means-for-you/ ）

EUROPEAN ARTIFICIAL INTELLIGENCE OFFICE

TIMELINE OF THE CODE OF PRACTICE DRAFTING PROCESS



(via https://digital-strategy.ec.europa.eu/en/news/ai-act-participate-drawing-first-general-purpose-ai-code-practice)

マルチステークホルダー プロセス

- EU AI Office は、行動規範の策定に参画するマルチステークホルダーを、 2024年8月25日まで、EU内外から公募（現在は受付終了） →日本からの参加者はごく少数
- 2024年9月30日、マルチステークホルダーが集まるキックオフ総会に、汎用目的AIモデルプロバイダー、下流プロバイダー、産業界、市民社会、学界、独立した専門家など約1,000人が参加
- 2025年4 月まで、3 回のオンライン会議が開催され、4つのワーキンググループ（WG）で議論しながら草案を作成
- 参加者は、各会議中または2週間以内にコメントを表明できる

Code Of Practice Plenary

WG 1
Transparency and copyright-related rules

Nuria Oliver
(Spain)
WG 1 Co-Chair
Transparency

Alexander Peukert
(Germany)
WG 1 Co-Chair
Copyright

- Vice Chair (Transparency): Rishi Bommasani (US)
- Vice Chair (Copyright): Céline Castets-Renard (France)

WG 2
Risk identification and assessment for systemic risk

Matthias Samwald
(Austria)
WG 2 Chair
Risk assessment

- Vice-Chair: Marta Ziosi (Italy)
- Vice-Chair: Alexander Zacherl (Germany)

WG 3
Technical risk mitigation for systemic risk

Yoshua Bengio
(Canada)
WG 3 Chair
Technical risk mitigation

- Vice-Chair: Daniel Privitera (Italy and Germany)
- Vice Chair: Nitarshan Rajkumar (Canada)

WG 4
Governance risk mitigation for systemic risk

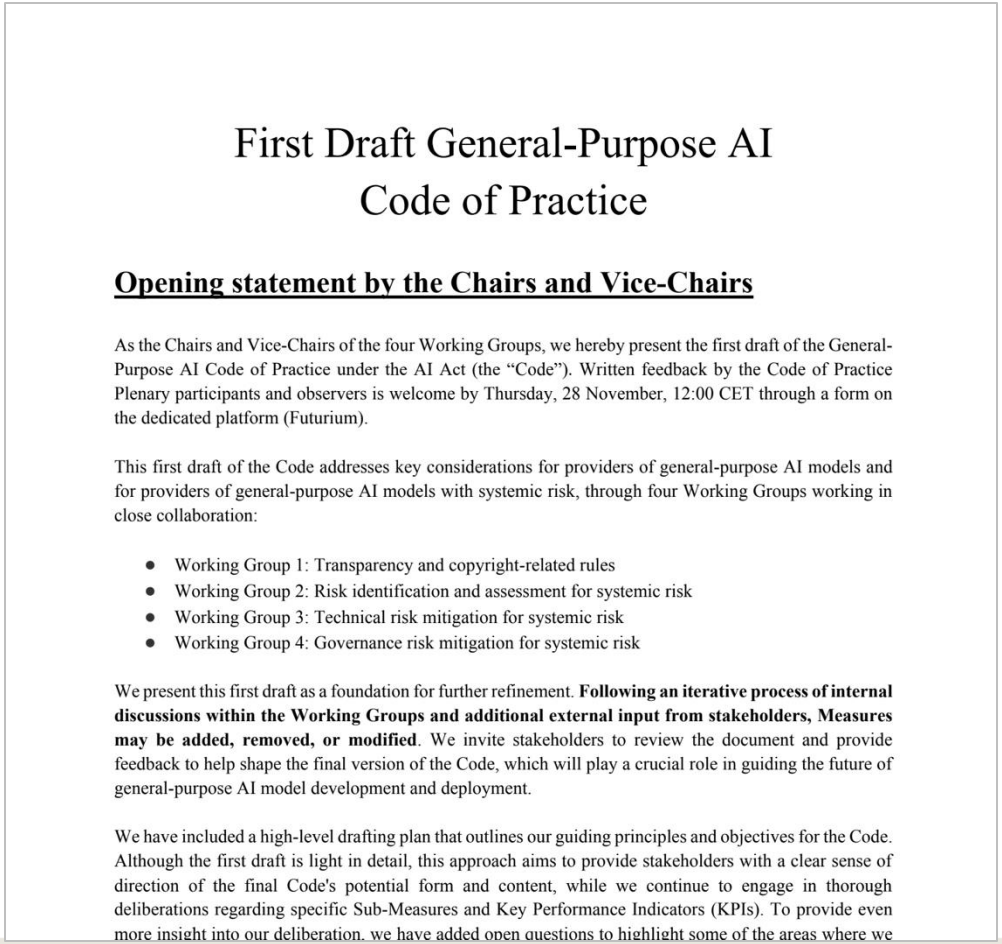
Marietje Schaake
(Netherlands)
WG 4 Chair
Internal Governance

- Vice Chair: Markus Anderljung (Sweden)
- Vice Chair: Anka Reuel (Germany)

（via https://digital-strategy.ec.europa.eu/en/news/meet-chairs-leading-development-first-general-purpose-ai-code-practice ）

現状

- 現在は、第1回会議と第2回会議の間
- 2024年11月14日に行動規範草案のファーストドラフト*とFAQ**が公表された
- 参加者はファーストドラフトに対するフィードバックを2024年11月28日締切で提出したところ



First Draft General-Purpose AI Code of Practice

Opening statement by the Chairs and Vice-Chairs

As the Chairs and Vice-Chairs of the four Working Groups, we hereby present the first draft of the General-Purpose AI Code of Practice under the AI Act (the "Code"). Written feedback by the Code of Practice Plenary participants and observers is welcome by Thursday, 28 November, 12:00 CET through a form on the dedicated platform (Futurium).

This first draft of the Code addresses key considerations for providers of general-purpose AI models and for providers of general-purpose AI models with systemic risk, through four Working Groups working in close collaboration:

- Working Group 1: Transparency and copyright-related rules
- Working Group 2: Risk identification and assessment for systemic risk
- Working Group 3: Technical risk mitigation for systemic risk
- Working Group 4: Governance risk mitigation for systemic risk

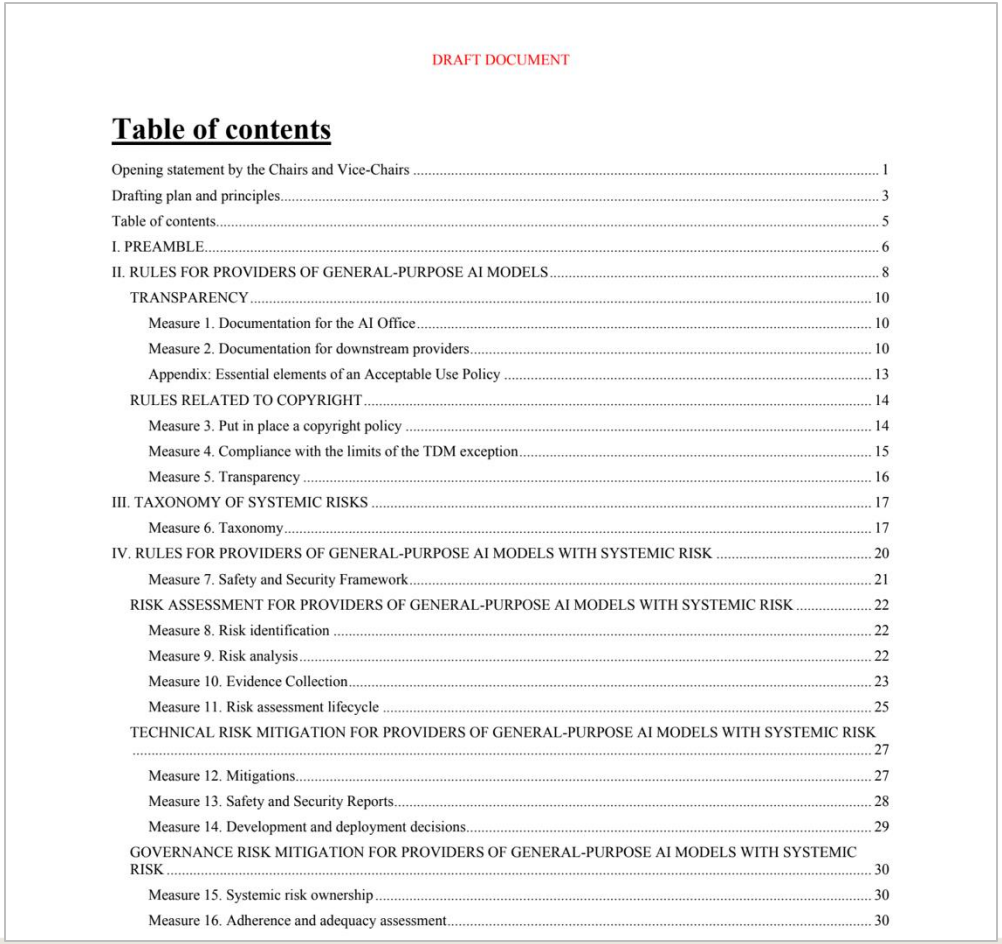
We present this first draft as a foundation for further refinement. **Following an iterative process of internal discussions within the Working Groups and additional external input from stakeholders, Measures may be added, removed, or modified**. We invite stakeholders to review the document and provide feedback to help shape the final version of the Code, which will play a crucial role in guiding the future of general-purpose AI model development and deployment.

We have included a high-level drafting plan that outlines our guiding principles and objectives for the Code. Although the first draft is light in detail, this approach aims to provide stakeholders with a clear sense of direction of the final Code's potential form and content, while we continue to engage in thorough deliberations regarding specific Sub-Measures and Key Performance Indicators (KPIs). To provide even more insight into our deliberation, we have added open questions to highlight some of the areas where we

* https://digital-strategy.ec.europa.eu/en/library/first-draft-general-purpose-ai-code-practice-published-written-independent-experts
** https://digital-strategy.ec.europa.eu/en/faqs/general-purpose-ai-models-ai-act-questions-answers

目次

- 行動規範には、目標（objective）、対策（measure）、KPIが含まれる
- 目標： EU AI法に準拠するもの
- 対策：プロバイダーが目標を達成するために実行する必要があるアクション群
  - サブ対策が記載されている場合も
  - 技術の進展を反映するためにレビュー期間中に調整される可能性
- KPI：目標が達成されたかどうかを評価するための指標を補足するもの



DRAFT DOCUMENT

Table of contents

目標：透明性

EU AI法　本文の引用

対策1:
EU AI Officeへの開示文書

対策2:
下流プロバイダーへの開示文書

開示情報の詳細を記載したリスト
（スライドでは省いてあるが
実際はとても長い）：
学習・検証・テスト用データに関する情報、アーキテクチャとパラメータ数、モデルと学習過程の設計仕様、学習時間や学習に用いた計算資源、推定消費電力など

TRANSPARENCY

LEGAL TEXT

Article 53(1)(a): "Providers of general-purpose AI models shall draw up and keep up-to-date the technical documentation of the model, including its training and testing process and the results of its evaluation, which shall contain, at a minimum, the information set out in Annex XI for the purpose of providing it, upon request, to the AI Office and the national competent authorities;"

Article 53(1)(b): "Providers of general-purpose AI models shall draw up, keep up-to-date and make available information and documentation to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems. Without prejudice to the need to observe and protect intellectual property rights and confidential business information or trade secrets in accordance with Union and national law, the information and documentation shall: (i), enable providers of AI systems to have a good understanding of the capabilities and limitations of the general-purpose AI model and to comply with their obligations pursuant to this Regulation; and (ii), contain, at a minimum, the elements set out in Annex XII;".

Measure 1. Documentation for the AI Office

Signatories commit to draw up and keep up-to-date the technical documentation of the model listed in the table below for the purpose of providing it, upon request, to the AI Office and the national competent authorities. The Signatories are encouraged to consider if the listed information can be disclosed, in whole or in part, to the public to advance public transparency.

Measure 2. Documentation for downstream providers

Signatories commit to draw up, keep up-to-date, and make available information and documentation listed in the table below to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems. The Signatories are encouraged to consider if the listed information can be disclosed, in whole or in part, to the public to advance public transparency.

AI Act reference	Detailing of information required	For the AI Office and national competent authorities	For downstream providers
Annex XI §1 1. and Annex XII 1.	General information: Signatories should detail general information about the provider of the general-purpose AI model and about the model itself to clearly identify and characterise the model, such as the model name, evidence of the provenance and authenticity of the model by means of e.g. a secure hash in the case binaries are distributed or TLS/SSL certificates in the case of a service, legal business name of the developer(s) and the owner(s) of the	✓	✓

（目標：ガバナンスリスク軽減）
対策18: 重大インシデントの報告

EU AI法　本文の引用

オープンクエッション：
起草者のChairたちからのお悩み相談

サブ対策18.1:
重大インシデントの報告プロセス

サブ対策18.2:
対応プロセスの準備

Measure 18. Serious incident reporting

LEGAL TEXT

Article 55(1)(c): “In addition to the obligations listed in Articles 53 and 54, providers of general-purpose AI models with systemic risk shall keep track of, document, and report, without undue delay, to the AI Office and, as appropriate, to national competent authorities, relevant information about serious incidents and possible corrective measures to address them;”.

Signatories commit to identify and keep track of serious incidents, as far as they originate from their general-purpose AI models with systemic risk, document and report, without undue delay, any relevant information and possible corrective measures to the AI Office and, as appropriate, to national competent authorities.

OPEN QUESTIONS

- What does a serious incident entail? Should the Code use the definition the AI Act uses for AI systems in Article 3(49) or is another definition more appropriate for general-purpose AI models with systemic risk?
- Under what conditions should a general-purpose AI model with systemic risk be judged to have indirectly led to a serious incident occurring?
- Are there suitable technical standards or best practices that can enable automated or streamlined reporting of serious incidents to the AI Office?

In order to satisfy Measure 18.:

Sub-Measure 18.1. Serious incident reporting processes

Signatories will set up processes (including by designating staff members) to identify, document, and report serious incidents and near-misses to the AI Office, as far as they originate from their general-purpose AI model with systemic risk.

Sub-Measure 18.2. Response readiness

Signatories will set up processes for responding to serious incidents, including pre-defining corrective measures that may be taken in response to serious incidents, along with an explanation of when they may be taken.

フィードバック提出時の雑感

- まだざっくりした案で全体的に詰め切れていない
  - Measureの記述密度が低い項目も多く、コメントしづらい
  - 例えば、 重大インシデントは主にアプリケーションレイヤーで生じると思うが、下流プロバイダー等の報告義務がどうなっているか、現案からは不明瞭など
  - 逆に言うと、域外適用を確保したいニーズなどがある場合、その提案ができる余地が大きい
- EU AI法の枠内であるはずだが、上乗せ／横出し規制になりかねない表記ぶり？
  - 例えば、条文上、情報開示の範囲は知的財産や営業秘密を考慮することになっているが（53条(1)(b)）、「Measure 2. Documentation for downstream providers」のみを読むと、そのような考慮がされているか若干怪しい
- （解釈上の疑義が明瞭になれば法学研究者としては満足だが）日本企業等のニーズがわからないと、日本の研究開発やビジネスにとって有益なフィードバックにならない
  - 削るべきところを削れるチャンスがあるが、自分だけの知見では活かしきれない

Thanks!

透明性についての規定（WG1関連）

- EU AI法の根拠　53条1項　GPAI提供者の義務
  a. EU-AI Officeおよび当局への技術情報の提供
  b. AIモデルをシステムに統合する提供者への情報提供

- CoPにおける記載
  - Measure 1：
    - EU-AI Officeおよび当局への技術情報の提供
    - **当該情報の一般公開を検討**
  - Measure 2
    - システム提供者に対する技術情報の提供

- 対象情報リスト

1. General information
2. Intended tasks and type and nature of AI systems in which it can be integrated
3. Acceptable use policies
4. Date of release and methods of distribution
5. Interaction of the model with external hardware or software
6. Versions of relevant software where applicable
7. Architecture and number of parameters
8. Modality and format of inputs and outputs
9. Licence
10. Technical means for integration into AI systems
11. Design specification and training process
12. Information on data used for training, testing and validation
13. Computational resources
14. Energy consumption
15. Testing process and results thereof

AIガバナンスについての規定（WG4関連）

- EU AI法の根拠　55条1項　システミックリスクのあるGPAI提供者の義務
  a. 敵対テストを含むモデル評価の実施
  b. リスク評価・軽減の実施
  c. 重大インシデントや是正措置について、EU-AI Officeおよび当局への情報提供
  d. サイバーセキュリティ対策の実施
- CoPにおける記載
  - Measure 8：システミックリスクの継続的特定
  - Measure 9：システミックリスクの分析
  - Measure 10：関連証拠の継続的収集
  - Measure 11：AIライフサイクルを通じたリスク評価と証拠収集
  - Measure 12：安全・セキュリティフレームワーク（SSF）へのリスク対策の明示
  - Measure 13：安全・セキュリティ報告書（SSR）の作成
  - Measure 14：リスクに対応したモデル導入決定システムの確立
  - Measure 15：システミックリスクの所有者（対応主体？）の決定
  - Measure 16：SSFの妥当性評価
  - Measure 17：独立した専門家によるシステミックリスクの評価
  - Measure 18：重大インシデントのEU-AI Officeおよび当局への報告
  - Measure 19：内部告発者保護
  - Measure 20：モデル性能に関するEU-AI Officeおよび当局への通知
  - Measure 21：CoP遵守状況に関する記録作成
  - Measure 22：SSFとSSRの一般公開

- 事業者にとって過重な負担にならないか？
- 第三者による検証は十分に実行できるのか？
- 消費者にとって役にたつ情報公開なのか？消費者は理解できるのか？

日本企業として考える必要があること

- CoPの規定はAI業界として過大なものとはなっていないか？
- CoPの規定は将来のイノベーションを阻害するものとはなっていないか？
- CoPの規定は日本のAI企業にとって過大な要求になっていないか？
- CoPの規定はSMEにとって過大な要求になっていないか？

- 問題があるならば
  - **CoP自体の変更による対処**
  - CoPの解釈による対処
  - EU当局への訴訟による対処
  - 日本政府を介した二国間、もしくは多国間交渉による対処

ご参加ありがとうございました
メールでアンケートを送らせていただきますので、
ご協力お願いいたします

問い合わせ先：
東京大学　江間研究室
tg-event@tc.u-tokyo.ac.jp