委員会
欧州委員会

ブリュッセル, 4.2.2025 C(2025)
884 final

附属書

以下についての

附属書

欧州委員会に対するコミュニケーション

欧州委員会からのコミュニケーション草案「規則（EU）2024/1689（AI 法）が定める禁止された人工知能行為に関する欧州委員会ガイドライン」の内容の承認

1. 背景と目的

(1) 2024 年 6 月 13 日付欧州議会・理事会規則（EU）2024/1689 は、人工知能に関する調和された規則を定め、特定の規則を改正するものである（「AI 法」）[1] 、2024 年 8 月 1 日に発効した。AI 法は、欧州連合における人工知能（以下「AI」）の上市、実用化、使用に関する調和された規則を定めたものである。[2]その目的は、民主主義や法の支配を含む連邦内の健康、安全、基本的権利の高水準の保護を確保しつつ、AI の技術革新と普及を促進することである。

(2) AI 法はリスクベースのアプローチに従い、AI システムを 4 つの異なるリスクカテゴリーに分類している：

(i) 容認できないリスク：基本的権利や連邦の価値観に受け入れがたいリスクをもたらす AI システムは、AI 法第 5 条により禁止されている。

(ii) リスクが高い：健康、安全、基本的権利に高いリスクをもたらす AI システムには、一連の要件と義務が課される。これらのシステムは、AI 法附属書 I および III とともに、AI 法第 6 条に従って「高リスク」に分類される。

(iii) 透明性リスク：限定的な透明性リスクをもたらす AI システムは、AI 法第 50 条に基づく透明性義務の対象となる。

(iv) リスクは最小かゼロか：リスクは極小から極小にとどまるが、プロバイダや展開者は自主的に行動規範を守ることができる。[3]

(3) AI 法第 96 条 1 項（b）に従い、欧州委員会は AI 法第 5 条で禁止されている慣行の実際の実施に関するガイドラインを採択することになっている。これらの禁止事項は、AI 法の発効から 6 ヵ月後、すなわち 2025 年 2 月 2 日から適用される。

(4) 本ガイドラインの目的は、法的明確性を高め、AI 法第 5 条の禁止事項の一貫した効果的かつ統一的な適用を確保するための欧州委員会の解釈に関する洞察を提供することである。本ガイドラインは、AI 法に基づく所轄当局の取締り活動や、AI システムのプロバイダや展開者が AI 法に基づく義務を遵守するための実践的な指針として役立つはずである。基本的権利と安全を保護するという AI 法の目的を達成しつつ、イノベーションを促進し、法的確実性を提供するために、禁止事項を適切な方法で解釈するよう努めている。

(5) 本ガイドラインは拘束力を持たない。AI 法の権威ある解釈は、最終的には欧州連合司法裁判所（以下、「CJEU」という。）

(6) 本ガイドラインのドラフトは、欧州委員会が主催した広範な協議の過程で収集された、AI システムのプロバイダや展開者、市民社会組織、学界、公的機関、企業団体など、さまざまな利害関係者からの

[1]人工知能に関する調和規則を定めた 2024 年 6 月 13 日付欧州議会および理事会規則（EU）2024/1689（人工知能法）（OJ L, 2024/1689, 12.7.2024）。

[2]第 1 条 AI 法

[3]第 95 条 AI 法

情報に基づいている。また、AI 委員会内の加盟国および欧州議会も協議を行った。本ガイドラインは、AI 法第 5 条の実践から得られた経験や、技術および市場の発展に照らして、定期的に見直される予定である。

(7) AI 法第 5 条の適用には、個々のケースで問題となる特定の状況を十分に考慮したケースバイケースのアセスメントが必要となる。従って、本ガイドラインに示された例は、単なる指標であり、個々のケースにおけるそのようなアセスメントの必要性を損なうものではない。

2. 禁止されている AI の慣行の概要

(8) AI 法第 5 条は、特定の AI システムを、操作的、搾取的、社会的統制的、監視的な行為のために EU 市場に上市、使用、または使用することを禁止している。AI 法第 28 条は、このような慣行は、人間の尊厳、自由、平等、民主主義、法の支配の尊重という EU の価値観や、非差別的権利（憲章第 21 条）、平等（憲章第 20 条）、データ保護（憲章第 8 条）、私的・家族的生活（憲章第 7 条）、子どもの権利（憲章第 24 条）など、欧州連合基本権憲章（「憲章」）に謳われている基本的権利に反するため、特に有害かつ濫用的であり、禁止されるべきであると明確にしている。また、AI 法第 5 条の禁止事項は、表現と情報の自由（憲章第 11 条）、集会と結社の自由（憲章第 12 条）、思想・良心・宗教の自由（憲章第 10 条）、効果的な救済と公正な裁判を受ける権利（憲章第 47 条）、無罪の推定と防御の権利（憲章第 48 条）を擁護することを目的としている。

2.1. AI 法第 5 条の禁止事項

(9) **禁止事項の概要**

提供	**禁止事項**	**内容**
第 5 条 (1)(a)	有害な操作、欺瞞	人の意識を超えたサブリミナル技法、または意図的に操作的もしくは欺瞞的な技法を展開し、行動を歪める目的もしくは効果を持つ AI システム。 、重大な危害を引き起こすか、または引き起こす可能性が合理的に高い。
第 5 条 (1)(b)	脆弱性の有害な悪用	年齢、障害、特定の社会的・経済的状況に起因する脆弱性を悪用し、行動を歪曲させ、重大な危害を引き起こす、または引き起こす可能性のある目的または効果を持つ AI システム。
第 5 条 (1)(c)	ソーシャル・スコアリング	社会的行動または個人的もしくは人格的特徴に基づいて自然人または集団を評価または分類する AI システムであって、社会的スコアが、無関係な社会的文脈からのデータである場合、またはそのような扱いが社会的行動に対して不当または不釣り合いである場合に、不利益または不利な扱いをもたらすもの。

第 5 条 (1)(d)	個々の犯罪リスクアセスメントと予測	プロファイリングまたは性格的特徴や特性のみに基づいて、人が犯罪を犯すリスクをアセスメントまたは予測する AI システム。ただし、犯罪行為に直接関連する客観的かつ検証可能な事実に基づく人間のアセスメントを支援する場合を除く。
第 5 条 (1)(e)	顔面を開発するための非標的スクレイピング 認識データベース	インターネットや閉回路テレビ（CCTV）映像から顔画像を無制限に収集し、顔認識データベースを作成または拡張する AI システム。
第 5 条 (1)(f)	感情認識	職場や教育機構において、医療上または安全上の理由がある場合を除き、感情を推測する AI システム
第 5 条 (1)(g)	バイオメトリック分類	人種、政治的意見、労働組合員、宗教的または哲学的信条、性生活または性的指向を推測または推論するために、生体データに基づいて人々を分類する AI システム。ただし、法執行の分野を含め、合法的に取得された生体データセットのラベリングまたはフィルタリングを除く。
第 5 条 (1)(h)	リアルタイムの遠隔生体認証 ('RBI')	ただし、特定の被害者の捜索、テロ攻撃を含む特定の脅威の防止、または特定の犯罪の容疑者の捜索のために必要な場合を除く（認可を含むさらなる手続き要件は、AI 法第 5 条 2 項 7 号に概説されている）。

2.2. 禁止事項の法的根拠

(10) AI 法は 2 つの法的根拠に支えられている：欧州連合機能条約（「TFEU」）第 114 条（域内市場法的根拠）と TFEU 第 16 条（データ保護法的根拠）である。TFEU 第 16 条は、法執行目的での遠隔生体認証（「RBI」）システム、法執行目的での生体分類システム、法執行目的での個人リスクアセスメントの使用禁止に関するパーソナルデータの処理に関する特定の規則の法的根拠となっている。[4]AI 法第 5 条に記載されたその他の禁止事項はすべて、TFEU 第 114 条に法的根拠がある。

[4]前文 3 AI 法。TFEU 第 16 条に基づく禁止に関して、アイルランドとデンマークには 2 つの関連するオプトアウトがある。TEU および TFEU に付属する「自由、安全および正義の分野における英国およびアイルランドの地位に関する議定書第 21 号（AFSJ）」に基づきアイルランドに認められている裁量権により、アイルランドは、法の執行を目的とした公共空間における RBI のリアルタイム使用の禁止に関する規則および同条に関連する手続き規則（AI 法第 5 条第 2 項から第 6 項）を適用しないことを決定することができる（前文 40 参照）。デンマークは TEU および TFEU の第 22 議定書を適用する際、オプトアウト協定の恩恵を受けており、TFEU 第 16 条に基づく禁止事項を完全に適用しないことを決定することができる（前文 41 参照）。

2.3. 重要な範囲：AI システムの「上市」、「使用開始」または「使用」に関する慣行

(11) AI 法第 5 条で禁止される行為は、特定の AI システムの上市、使用開始、または使用に関するものである。[5]リアルタイムの遠隔生体認証（「RBI」）システムに関しては、AI 法第 5 条(1)（h）の禁止はその使用のみに適用される。AI 法第 3 条第 1 項は、AI システムを構成するものを定義している。AI システムの定義に関するガイドラインは、その定義に関する欧州委員会の解釈を示している。

(12) AI 法第 3 条 9 項によれば、AI システムの上市とは、「AI システムを［...］連合市場で初めて入手可能にすること」である。入手可能」とは、「商業活動の過程において、有償であるか無償であるかを問わず、連合市場で頒布または使用するために」システムを供給することと定義される。[6]AI システムの利用可能化は、アプリケーション・プログラミング（API）を通じたシステムやそのサービスへのアクセス、クラウド経由、直接ダウンロード、物理的なコピー、物理的な製品への組み込みなど、供給手段に関係なく対象となる。

> 例えば、第三国のプロバイダが域外で開発した RBI システムは、1 つまたは複数の加盟国で支払いと引き換えに、または無償で提供されることにより、初めて上市される。このような上市は、API またはその他のユーザーインターフェイスを通じてオンライン上でシステムへのアクセスを提供することによって行われる。

(13) AI 法第 3 条(11)は、「AI システムを、その意図された目的のために、導入者に最初に使用させるため、または組合内で自ら使用するために供給すること」と定義しており、したがって、サードパーティへの最初に使用させるための供給も、自社での開発・展開も対象となる。システムの意図された目的とは、「プロバイダが AI システムを意図する用途であり、使用説明書、販促用資料、販売用資料、明細書、および技術文書においてプロバイダが提供する情報に明記されている具体的な使用状況や条件を含む[7]。

> 例えば、プロバイダが加盟国外で RBI システムを構築し、そのシステムを法執行当局または加盟国に設立された民間企業に供給して初めて使用されることにより、サービスが開始される。
>
> 例えば、ある認可機関がスコアリング・システムを自社開発し、家計手当受給者の不正リスクを予測するために展開する。

(14) AI システムの「**使用」は**、AI 法では明確に定義されていないが、上市またはサービス開始後のライフサイクルのあらゆる時点におけるシステムの使用または展開を含むと広義に理解すべきである。また、より複雑なシステム、プロセス、インフラの一部としてなど、AI システムを利用する者のサービスやプロセスにおける AI システムの統合も含まれる。AI システムのプロバイダは、AI システムを上市する前に、合理的に予見可能な使用条件（意図された使用及び合理的に予見可能な誤用[8]）を考慮しなければなら

[5]これらの用語の定義については、欧州委員会告示「EU 製品規則 2022 の実施に関する『ブルーガイド』」2022/C 247/01 の第 2 節も参照のこと

[6]AI 法第 3 条 10 項

[7]AI 法第 3 条 12 項

[8]AI 法第 3 条 12 項および 13 項参照。

ないが、展開者は、システムの使用に関する合法的条件を考慮する責任を負う。[9]AI 法第 5 条の目的上、「使用」への言及は、禁止行為に相当する可能性のある AI システムの誤用（「合理的に予見可能」か否かを問わない）を含むと理解すべきである。[10]。

例えば、使用者がワークスペースでの感情を推測するために使用する AI システムは、医療や安全の目的で使用される場合を除き、禁止されている（AI 法第 5 条(1)（f））。この禁止は、プロバイダ（システムの供給者）が使用者（雇用主）との契約関係、すなわち利用規約においてそのような使用を除外しているか否かにかかわらず、展開者に適用される。

2.4. 個人の範囲：責任ある行為者

(15) AI 法では、AI システムに関する事業者をプロバイダ、展開事業者、輸入事業者、頒布事業者、製品製造者とカテゴリー分けしている。本ガイドラインでは、AI 法第 5 条で禁止されている行為の範囲を考慮し、プロバイダと展開者のみに焦点を当てる。

(16) AI 法第 3 条第 3 項によれば、プロバイダとは、AI システムを開発し、または開発させ、それを連合域内の市場に流通させ、もしくは自己の名称もしくは商標[11]（上記第 2.3 項参照）の下でサービスを提供する自然人もしくは法人、公的機関、代理店その他の団体である。域外に設立され又は域外に所在するプロバイダは、当該システムを域内で上市又は使用する場合（[12]）、又は AI システムの出力が域内で使用される場合（[13]）、AI 法の規定に従う。プロバイダは、AI システムを上市または使用開始する前に、その AI システムがすべての関連要件を満たしていることを確認しなければならない。

例えば、RBI システムのプロバイダとは、その商標の下で組合内でシステムを販売するシステムの製造事業者である。このようなシステムのプロバイダは、自社でシステムを開発し、自社で使用するためにサービスを提供する公的機関である可能性もある。

(17) **展開**者とは、自然人または法人、公的機関、代理店、その他の団体であり、その権限の下で AI システムを使用する。[14]AI システムに対する「認可」とは、システムの展開決定と実際の使用方法に対する責任を負うことと理解すべきである。展開者は、その事業所または所在地が連邦（[15]）内にある場

[9]これらの用語の定義については、欧州委員会告示-EU 製品規則 2022 の実施に関する「ブルーガイド」、2022/C 247/01、2.8 項も参照のこと。

[10]前文 28 AI 法。

[11]AI 法第 3 条第 3 項、第 9 項、第 11 項。高リスク AI システムに関して、AI 法第 25 条は、以下を想定している。1.頒布事業者、輸入事業者、展開事業者又はその他のサードパーティは、本規則の適用上、高リスク AI システムのプロバイダとみなされ、以下のいずれかの場合、第 16 条に基づくプロバイダの義務を負う：(b) 既に上市され又は既に稼働している高リスク AI システムに対して，第 6 条に基づき高リスク AI システムであることを維持するような実質的な改変を行う場合；(c) 高リスクと分類されておらず、既に上市又は使用開始されている汎用 AI システムを含む AI システムの意図された目的を、当該 AI システムが第 6 条に従って高リスク AI システムとなるように変更する場合。

[12]AI 法第 2 条第 1 項（a）。

[13]AI 法第 2 条第 1 項（c）。

[14]AI 法第 3 条第 4 項

[15]AI 法第 2 条第 1 項第 2 号。

合、または第三国に所在する場合、AI システムの出力が連邦（[16] ）内で使用される場合、AI 法の適用範囲に入る。

(18) AI システムの展開者が、そのシステムが使用される権限を有する法人、すなわち法執行当局や民 間警備会社である場合、その法人の手順の範囲内で、その法人の管理の下で行動する 個々の従業員は、展開者とみなされるべきではない。法人が、その責任と管理の下で、その法人に代わってシステムの運用にサードパーティ （請負業者、外部スタッフなど）を関与させる場合も、展開者にとどまる。

(19) 事業者は、AI システムに関して複数の役割を同時に果たすことができる。例えば、ある事業者が独自の AI システムを開発し、その後それを使用する場合、そのシステムが有償または無償で提供された他の事業者によっても使用されるとしても、その事業者はそのシステムのプロバイダであると同時に展開者であるとみなされる。

(20) AI のライフサイクルの全段階において、AI 法の継続的な遵守が求められる。このため、AI システムがそのライフサイクルを通じて AI 法に準拠し続け、AI 法第 5 条で禁止される行為に至らないよう、上市された AI システム、または連合内で使用される AI システムの継続的な監視と更新が必要となる。AI システムのプロバイダと展開者は、その役割と、禁止行為を回避するためのシステムの設計、開発、実際の使用に対する管理に応じて、異なる責任を負う。各禁止事項について、バリューチェーンの誰が具体的な予防・緩和措置を採用し、AI 法の目的及び考え方に沿った AI システムのコンプライアンスな開発及び使用を確保するのに最も適した立場にあるかを考慮し、これらの役割と責任を比例的に解釈すべきである。

2.5. AI 法の適用除外

(21) AI 法第 2 条は、AI 法第 5 条に記載された禁止事項の実際の適用を完全に理解するために関連する、いくつかの一般的な適用除外をプロバイダとして規定している。

2.5.1. 国家安全保障、防衛、軍事目的

(22) AI 法 2 条 3 項によれば、AI 法は EU 法の適用範囲外の領域には適用されず、加盟国が国家安全保障に関する事業体をどのような種類で委託しているかにかかわらず、いかなる場合にも加盟国の国家安全保障に関する権限に影響を及ぼすべきではない。AI 法は、「軍事、防衛、国家安全保障の目的のためにのみ、上市され、使用され、改造の有無にかかわらず使用される」AI システムを、それらの活動を実施する事業体の種類にかかわらず、その範囲から明確に除外している。したがって、この除外が適用されるかどうかは、AI システムの目的や用途に依存するのであって、そのシステムを使って活動を行う事業体に依存するわけではない。

(23) CJEU によれば、**「国家安全保障」とは**、「国家の本質的機能と社会の基本的利益を守ることを第一義とし、 、国の基本的な憲法的、政治的、経済的、社会的構造を著しく不安定化させる可能性のある活動、特にテロ活動のように社会、国民、国家そのものを直接脅かす可能性のある活動の防止と処

[16] AI 法第 2 条第 1 項（c）。

罰を含む」ものである。[17]国家安全保障は、例えば交通安全、[18] 、司法の組織や運営に関連する活動には適用されない[19]。CJEU が述べているように、「自国の本質的な安全保障上の利益を定義し、自国の内外の安全保障を確保するために適切な措置を採用するのは加盟国である（中略）国家安全保障を保護する目的でとられた国内措置（中略）により、EU 法が適用できなくなったり、加盟国が同法を遵守する義務が免除されたりすることはありえない[20] 。

(24) AI 法第 2 条(3)項第 2 号の除外が適用されるためには、AI システムは、専ら軍事、 防衛又は国家安全保障の目的のために上市、稼働又は使用されなければならない。前文 24 AI 法はさらに、「**専ら」という**概念をどのように解釈すべきか、また、そのような目的で使用される AI システムが AI 法の適用範囲に含まれるのはどのような場合かを明確にしている。

> 例えば、上市され、運用が開始され、軍事・防衛・国家安全保障の目的で使用されている AI システムが、民間・人道目的、法執行・公安目的など、他の目的で（一時的または恒久的に）使用される場合、そのシステムは AI 法の適用範囲に含まれることになる。この場合、AI システムを他の目的に使用する事業体は、AI システムが AI 法に適合していることを確認しなければならない。

(25) さらに、AI 法 24 条は、軍事、防衛、国家安全保障という除外された目的のために上市または使用開始された AI システムと、民間や法執行目的など 1 つ以上の除外されない目的のために上市または使用開始された AI システム（いわゆる「**デュアルユース」**システム）は、AI 法の適用範囲に含まれることを明確にしている。これらのシステムのプロバイダは、AI 法の要件に準拠していることを確認する必要がある。

> 例えば、ある企業が法執行や国家安全保障を含む様々な目的のために RBI システムを提供する場合、その企業は「デュアルユース」システムのプロバイダであり、AI 法の要件への準拠を保証しなければならない。

(26) しかし、AI システムが AI 法の適用範囲に入る可能性があるという事実は、国家安全保障、防衛、軍事 の活動を行う事業体が、それらの活動を行う事業体の種類にかかわらず、国家安全保障、軍事、防衛の目的で当該システムを使用する能力に影響を及ぼすべきではない[21] 。

> 例えば、国家安全保障局または民間事業者が、国家情報機関から国家安全保障目的（情報収集など）でリアルタイム RBI システムを使用するよう命じられた場合、そのような使用は AI 法の適用範囲から除外される。

---

[17]2020 年 10 月 6 日付司法裁判所判決（*La Quadrature du Net and Others*, C-511/18, C-512/18 and C-520/18,EU:C:2020:791,パラグラフ 135）；2023 年 6 月 5 日付司法裁判所判決（*Commission* v *Poland*, C-204/21, EU:C:2023:442,パラグラフ 318）。

[18]2021 年 6 月 22 日付ラトビア共和国司法裁判所判決、C-439/19、EU:C:2021:504、パラグラフ 68。

[19]2023 年 6 月 5 日、欧州司法裁判所の判決（*Commission v Poland*, C-204/21, EU:C:2023:442, paragraph 319）。

[20]2020 年 10 月 6 日の司法裁判所の判決、プライバシー・インターナショナル、C-623/17、EU:C:2020:790、パラグラフ 44。

[21]前文 24 AI 法。

(27) 国家安全保障の適用除外を明確に定義することは、AI システムが上市され、使用され、AI 法の適用範囲に含まれる法執行目的に使用される場合に特に重要である。これは、AI 法第 5 条(1)（d）と（h）にそれぞれ規定されている、個別の犯罪予測とアセスメントに関する禁止事項、および法執行目的でのリアルタイム RBI システムの使用に関する禁止事項に関連している。警察その他の法執行当局は、公共の安全に対する脅威を保護・防止することを含め、犯罪の予防、検知、捜査、訴追、または刑事罰の執行を任務としている[22] 。AI システムがこのような目的で使用される場合は、必ず AI 法の適用範囲に入る。

(28) 欧州刑事警察機構（Europol）や、フロンテックス（Frontex）などの欧州連合（EU）の治安機関の活動は、AI 法の適用範囲に含まれる。

2.5.2. 第三国との司法・法執行協力

(29) AI 法 2 条 4 項によれば，AI 法は，第三国の公的機関又は国際機関には適用されない。ただし，これらの機関又は組織が，連邦又は 1 以上の加盟国との法執行及び司法協力のための国際協力又は協定の枠組みにおいて AI システムを使用する場合には，当該第三国又は国際機関が個人の基本的権利及び自由の保護に関して適切な保護措置を提供することを条件とする。関連する場合、この除外は、当該法執行及び司法協力を支援する特定の業務を遂行するよう当該第三国から委託された民間事業体の活動を対象とすることができる。[23]同時に、この除外が適用されるためには、これらの協力の枠組み又は国際協定は、法執行及び司法の領域で使用される AI システムの監督について権限を有する市場監視当局が評価する、個人の基本的権利及び自由の保護に関する適切なセーフガードを含んでいなければならない。[24]前文 22 AI 法は、取得者である国家当局と、そのような AI のアウトプットを連邦内で利用する連邦の機構、団体、事務所、機関が、その利用が連邦法に準拠していることを保証する責任を負うことを明確にしている。将来、これらの国際協定が改定され、または新たな協定が締結される場合、 締約国は、これらの協定を AI 法の要件に合わせるよう最大限の努力をすべきである。

2.5.3. 研究開発

(30) AI 法第 2 条 8 項によれば、AI 法は「上市または実用化される前の AI システムまたは AI モデルに関する研究、試験または開発活動」には適用されない。この適用除外は、AI システムが上市または実用化された時点で適用される AI 法の市場主義の論理に沿ったものである。

例えば、研究開発（R&D）の段階では、AI 開発者は、消費者向けのアプリケーションで使用された場合、操作的とみなされ、AI 法第 5 条(1)（a）の対象となり得る技術を含む可能性のある新しい機能を実験・テストする自由を有する。AI 法は、AI 技術を洗練させ、上市前に安全性と倫理基準を

[22]AI 法第 3 条 46 項

[23]AI 法の前文 22 を参照のこと。

[24]前文 22 および AI 法 74 条 8 項を参照のこと。

満たすようにするためには、初期段階の研究開発が不可欠であると認識し、このような実験を認めている。

(31) AI 法第 25 条で明確にされているように、AI 法はイノベーションを支援することを目的としており、AI 技術を進歩させ、科学の進歩とイノベーションに貢献する科学研究の重要性を認識している。そのため、AI 法第 2 条第 6 項では、「科学的研究及び開発のみを目的として特別に開発され、実用化された AI システム又は AI モデル（その出力を含む）」の除外を規定している。

例えば、AI によるサブリミナル刺激や欺瞞的刺激に対する認知・行動反応の研究は、人間と AI の相互作用に関する貴重な洞察を提供し、将来のより安全で効果的な AI アプリケーションに情報を提供することができる。このような研究は、AI 法第 5 条(1)（a）の禁止にかかわらず、AI 法の適用範囲から除外されているため、認められている。

(32) しかし、AIA 法第 2 条第 8 項の除外は、AI システムがそのような研究開発活動の結果として上市され たり、使用されたりする場合には、AI 法を遵守する義務を損なうものではない□[25]□AI 法（[26] ）の意味における実環境での試験も、この除外の対象とはならない。

例えば、ある自治体がカーニバルの期間中、街頭で RBI システムを使用した顔認識ソフトウエアのテストを希望する場合、実世界の状況下でシステムによって識別されるボランティアを募集する。実世界でのテストは AI 法第 2 条第 8 項の適用除外に該当しないため、システムが AI 規制のサンドボックスでテストされるか、あるいは AI 法第 60 条および第 61 条に規定されるサンドボックス外の実世界条件でのテストのための特別制度に従わない限り、計画されたテストは AI 法の RBI システムに対する の要件に完全に準拠しなければならない。[27]

(33) いずれにせよ、研究開発活動（AI 法の適用範囲から除外される場合を含む）は、科学研究に関する公認の倫理的・専門的基準に従って実施されるべきであり、適用される EU 法[28] （例えば、引き続き適用されるデータ保護法）に従って実施されるべきである。

2.5.4. 個人的な非専門活動

(34) AI 法 2 条 10 項は、AI 法は「純粋に個人的な非専門的活動の過程でシステムを使用する自然人であるプロバイダの義務には適用されない」と規定している。展開者の定義には、このような活動に従事するユーザーも含まれない（上記 2.4.項参照）。自然人が定期的に経済的利益を得る活動、または専門的、事業的、貿易的、職業的、またはフ リーランスの活動に関与する活動は、「専門的」活動とみなされるべきである。個人的な」という指定は、非専門家という修飾語であり、その人が個人的な立

[25]前文 25 AI 法。

[26]AI 法第 3 条第 57 項によれば、「実環境における試験」とは、信頼できる確実なデータを収集し、AI システムが本規則の要求事項に適合していることをアセスメントし検証することを目的として、実験室またはその他の模擬環境以外の実環境において、AI システムをその意図された目的のために一時的に試験することを意味する。AI 法は、このような実環境での試験について特別な制度を定めており、試験に参加する者の自由意思に基づくインフォームド・コンセントの取得など、第 57 条または第 60 条に定める条件がすべて満たされていれば、本規則の意味における AI システムの上市または使用開始には該当しない（AI 法第 60 条参照）。

[27]AI 法には、AI 規制サンドボックスと実世界テストに関する詳細かつ具体的な義務が含まれている。AI 法第 57 条参照のこと

[28]前文 25 AI 法。

場と非専門的な立 場の両方で行動することを意味する。従って、例えば、犯罪活動は純粋に個人的な活動とはみなされないため、この除外は犯罪活動を包含すべきではない。

例えば、自宅で顔認識システムを使用している認可者は（例えば、アクセス管理や自宅入口の安全監視のために）、AI 法第 2 条 10 項の適用除外に該当するため、たとえ映像（の一部）を法執行当局に送信する必要がある場合でも、AI 法に基づく展開者の義務の対象とはならない。

これとは対照的に、フリーランス、ジャーナリスト、医師などの職業的活動のために AI システムを使用する自然人は、AI 法に基づく顔認識システムの展開者に対する義務を遵守する必要がある。自然人が、専門的な立場で行動する展開者の代理または認可の下で行動する場合の使用も、AI 法の適用範囲に入る。

さらに、犯罪行為は、たとえ経済的利益の追求や獲得がなくても、純粋に個人的な活動とはみなされない。消費者保護法やデータ保護法、国内行政法への不遵守など、その他の違法行為については、AI 法の除外規定が適用されるが、その他の関連する法的枠組みは引き続き適用される）。

(35) AI 法第 2 条(10)の除外は、システムを純粋に個人的な非専門的活動に使用する場合の 展開者の義務に関してのみ適用される。システムを上市又は使用するプロバイダ、その他の専門的な配備者、及び輸入事業者や頒布事業者のようなその他の責任ある事業者の義務に関しては、そのようなシステムは依然として AI 法の範囲内にある。

例えば、感情認識システムは、自然人が純粋に個人的な非専門的活動に使用することを意図している場合、AI 法第 6 条に分類される高リスクの AI システムであることに変わりはなく、AI 法を完全に遵守しなければならない。同時に、純粋に個人的な非専門的使用のためにそれを使用する展開者（例えば自閉症者）は、AI 法に基づく展開者に対する特定の義務の対象とはならず、その使用は範囲外となる。

2.5.5. フリーおよび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

(36) AI 法第 2 条第 12 項によれば、AI 法は、フリー・オープンソース・ライセンス[29] の下でリリースされた AI システムには適用されない。ただし、高リスクの AI システムとして、または AI 法第 5 条（禁止される AI 慣行）もしくは第 50 条（特定の AI システムに対する透明性義務）に該当する AI システムとして、上市または使用開始される場合は例外である。つまり、AI システムのプロバイダは、上市またはサービスインした AI システムが AI 法第 5 条の禁止行為に該当する場合、この除外措置の恩恵を受けることができない。

[29]前文 102 AI 法は、フリーでオープンソースソフトウェアのライセンスに基づくソフトウェアやデータのリリースについて、「オープンに共有され、ユーザーが自由にアクセスし、使用し、修正し、再配布することができる」と説明している。

2.6. 禁止事項と高リスク AI システムの要件の相互関係

(37) AI 法第 5 条が禁止する AI の慣行は、AI 法第 6 条に従ってハイリスクに分類された AI システム、特に附属書 III に列挙された AI システムとの関連において考慮されるべきである。[30]というのも、ハイリスクに分類される AI システムの使用は、場合によっては、AI 法第 5 条の禁止事項の 1 つ以上の条件がすべて満たされた場合、特定のケースにおいて禁止行為として認められる可能性があるからである。逆に、AI 法第 5 条に記載された禁止事項の例外に該当する AI システムのほとんどは、ハイリスクと認定される。

例えば、感情認識システムは、AI 法第 5 条(1)(f)の禁止条件を満たさない場合、AI 法第 6 条第 2 項および附属書 III のポイント(1)(c)に従い、高リスク AI システムに分類される。同様に、信用スコアリングや健康保険・生命保険におけるリスクアセスメントに使用されるような AI ベースのスコアリングシステムも、AI 法第 5 条(1)(c)の禁止条件を満たさない場合、高リスク AI システムとみなされる。[31]もう一つの例は、 、医療サービスや社会保障給付など、ハイリスクに分類されるような、人を評価し、必要不可欠な公的扶助給付やサービスを受ける権利があるかどうかを判断する AI システムである。[32]このようなシステムが容認できない社会的評価を伴い、AI 法第 5 条(1)(c)の条件を満たす場合、その上市、使用開始、使用は連合内で禁止される。

このような場合、プロバイダがリスクアセスメントとマネジメントを行い、高リスク AI システムに関するその他の要件（データガバナンス、透明性、人的監視など）を遵守すること、また、使用説明書に従って適切に使用すること、人的監視（第 26 条）、場合によっては基本的権利の影響評価（第 27 条）を行うことが、上市または展開された高リスク AI システムが合法的であり、禁止行為に当たらないことを保証するのに役立つはずである。

(38) 最後に、附属書 III の高リスクのユースケースに該当するにもかかわらず、AI 法第 6 条第 3 項に基づいて例外的に高リスクとみなされない AI システムも、AI 法第 5 条の禁止事項の範囲に含まれる可能性がある。AI 法第 6 条第 3 項は、AI システムをハイリスクではないとみなす結果をもたらすだけであり、そのような AI システムを AI 法や禁止事項の範囲から排除するものではない。

2.7. 汎用 AI システムおよび目的別システムへの禁止事項の適用

(39) この禁止事項は、「意図された目的」[33] 、「汎用」（すなわち、様々な目的に対応可能）であるか否かを問わず、直接使用するため、または他の AI システムに統合するための、あらゆる AI システムに適用される[34]したがって、各事業者は、バリューチェーンにおけるシステムの役割と管理に基づいて、責任あ

[30]このリストでは、バイオメトリクスに基づく AI システムだけでなく、雇用、教育、公共・民間サービスへのアクセス、法執行など、特定の領域で特定の目的に使用される AI システムも対象としている。

[31]これは前文 58 と附属書 III に明記されている。

[32]前文 58 AI 法。

[33]AI 法第 3 条第 12 項において、AI システムがプロバイダによって意図される用途として定義され、使用説明書、販売促進資料、明細書、および技術文書においてプロバイダによって提供される情報において仕様される、具体的な使用状況および条件を含む。

[34]AI 法第 3 条 66 項参照。

る安全な AI システムの提供と利用を確保するために、AI 法の 2 つの目的を達成する観点から、そのリスクと便益のバランスをとりながら、最適な措置を講じるべきである。

(40) したがって、展開者は、システムのプロバイダが実装する安全ガードレールを回避しないことを含め、AI 法第 5 条で禁止されている方法で AI システムを使用しないことが期待されている。弊害は多くの場合、AI システムの実際の使われ方から生じるが、プロバイダには、汎用 AI システムを含め、AI 法第 5 条で禁止されるような動作をしたり、直接使われたりする可能性が合理的に高い AI システムを上市したり、利用したりしない責任もある。[35]この文脈において、プロバイダは、セーフガードを構築し、そのような有害な行動や誤用を防止・緩和するための効果的かつ検証可能な措置を、それらが合理的に予見可能であり、かつ、その措置が特定の AI システムや事案の状況に応じて実行可能かつ比例的である（ ）範囲で講じることも期待される。プロバイダは、展開者との契約関係（すなわち、AI システムの使用条件）において、禁止された行為に対する AI システムの使用を排除し、展開者に対する使用説明書や必要な人的監視に関する適切な情報を提供することも期待される。

> 例えば、チャットボットとして使用される汎用の AI システムは、重大な危害を引き起こす可能性の高い、操作的で欺瞞的な手法を展開する可能性がある。AI 法第 5 条(1)第 1 号（a）に基づき、AI システムの禁止行為や、操作・欺瞞・重大な危害を引き起こす可能性が合理的に高い利用を防止するため、プロバイダは適切かつ相応の措置（例．適切な安全かつ倫理的な設計、技術的及びその他のセーフガードの統合、使用の制限、透明性及び利用者の管理、使用説明書における適切な情報）、チャットボットが利用者、その他の者又は集団に重大な危害を与えないことを保証するために、AI システムが上市される前に（AI 法第 5 条(1)第 1 号）。

(41) 特定の場合、特に禁止事項がシステムの極めて具体的な目的（[36] ）と関連 している場合、プロバイダは他の予防策や緩和策を統合する可能性が限られる可能性があり、主に、 展開者と必要な人的監視に対する適切な指示と情報の提供、およびシステムの禁止された使用の制限に頼らざるを 得ない。適切な場合、そのような措置には、AI システムが供給される手段や、誤用の可能性についてプロバイダが自由に利用できる情報によっては、その制限が遵守されているかどうかを監視することも含まれる。不正使用を検知するための可能な監視措置は、展開者の活動を一般的に監視するようなものであってはならず、EU 法に沿ったものでなければならない。

> 例えば、医療や安全上の理由による例外が適用されない限り、感情を認識または推論できる汎用 AI システムは、職場や教育機構で展開者によって使用されるべきではない。しかし、プロバイダは、システムの感情認識機能が使用される具体的な状況や、AI 法第 5 条(1)（f）の禁止の例外が適用されるかどうかを知る立場にないかもしれない。しかしながら、そのようなプロバイダは、利用規約においてそのような禁止された使用を明示的に除外することができ、また、展開者を導くための使用説明書に適切な情報を含めることができる。また、プロバイダは、システムが特定の展開者によってこのような特定の禁止目的のために誤用されていることを認識した場合、適切な措置を講じることが期待される。例えば、この

[35]これは特に、第 5 条(1)（h）のリアルタイム RBI システムの禁止が使用のみに適用されることを除いて、AI 法第 5 条に列挙されたすべての禁止事項において「上市」または「使用開始」に言及していることから導かれる。

[36]AI 法第 5 条(1)(d)-(h)。

ような誤用が報告された場合、またはプロバイダが他の方法で認識した場合（システムがプロバイダの管理下にあるプラットフォームを通じて直接運用され、プロバイダがチェックを行う場合など）である。

2.8. 禁止事項と他の EU 法との相互関係

(42) AI 法は、他の連邦法、特に基本的権利の保護、消費者保護、雇用、労働者の保護、製品安全[37]を損なうことなく、あらゆる分野に水平適用される規制である。AI 法は、予防と安全の論理（AI システムを上市したり、特定の方法で使用したりしてはならない）を通じてこうした法律を補完し、他の法律では禁止されていないような特定の有害な AI 慣行を取り上げることで、さらなる保護を提供する。さらに、AI システムのライフサイクルの初期段階（すなわち、上市とサービス開始）と展開（すなわち、使用）に対処することで、AI 法の禁止事項は、AI のバリューチェーンの様々な時点で、AI に関わる有害な行為に対して措置を講じることを可能にする。

(43) 同時に、AI 法は、AI の慣行が他の同盟法（[38] ）に該当する場合に適用される禁止事項には影響しない。したがって、AI システムが AI 法で禁止されていない場合でも、その使用は他の第一次または第二次連邦法に基づいて禁止または違法となる可能性がある（例えば、データ保護法で要求されるパーソナルデータの処理に法的根拠がない、連邦法で禁止される識別的差別があるなど、特定のケースで基本的権利が尊重されないため）。したがって、AI 法の禁止事項の遵守は、AI システムのプロバイダや展開者に適用される他の連邦法を遵守するための十分な条件とはならない。

例えば、職場で使用される AI 対応感情認識システムは、医療または安全のために使用されるため、AI 法第 5 条(1)(f)の禁止事項から免除されるが、データ保護法、労働安全衛生を含む雇用および労働条件に関する連邦法および国内法の対象者である。[39]

(44) AI システムの上市や使用に関連する特定の活動が、他の連邦法でもカバーされている場合、AI 法は、異なる規定の一貫した実施を保証することを目的としている。さらに、AI 法の施行に責任を負う管轄当局と、AI 法第 77 条および AI 法のその他の規定に従って基本的権利を保護する当局との間の効果的な協力を可能にする。より一般的には、TEU 第 4 条第 3 項に従い、関係諸機関は、連邦法に基づくそれぞれの任務を遂行する際に、誠実に協力する義務がある。

(45) AI システムはしばしば、個人を特定できる自然人に関する情報（「パーソナルデータ」）を処理するためである。[40]禁止事項や文脈にもよるが、このようなシステムに関連する最も関連性の高い法律行為は、個人データの処理に関する自然人の保護およびそのようなデータの自由な移動に関する規則（EU 一般データ保護規則、以下「GDPR」）、予防を目的とする権限のある当局による個人データの処理に関する自然人の保護に関する指令（EU）2016/680 である、EU 機関、団体、事務所および機関に対するデータ保護規則を定めた規則（EU）2018/1725（以下「EUDPR」という。AI 法第 2 条第 7 項に従い、これらの法律は影響を受けず、EU データ保護法と整合的かつ補完的な

[37]AI 法第 2 条と前文 9。

[38]AI 法第 5 条 8 項

[39]AI 法前文 9 も参照のこと。

[40]AI 法 2 条 7 項、AI 法前文 10 項も参照のこと。

AI 法とともに引き続き適用される。これらの EU データ保護規則のいくつかの側面は CJEU によって明確化されており、欧州データ保護委員会は一連のガイドラインを採択している（例えば、「プロファイリング」の概念について[41] 、同じ概念を使用していることから、AI 法第 5 条(1)（d）の禁止に特に関連している）。

(46) 法執行目的での生体データ分類システムおよびリアルタイム RBI システムの使用に関する禁止／制限につい ては、AI 法が第 10 条 LED の特別法として適用されるため、そのような使用および関係する生体データの処理 が網羅的に規制される[42] 。この文脈において、AI 法は指令（EU）2016/680 の第 8 条に基づくパーソナルデータの処理の法的根拠を提供することを意図していない。同指令の他のすべての規定は、AI 法に規定された条件に加えて、特に、AI 法第 5 条(1)(h)の限定的な例外を条件として、許可された場合の法執行目的でのリアルタイム(RBI)システムの使用に適用される。より一般的には、管轄法執行当局（すなわち、LED 第 3 条第 7 項に基づく管轄当局）が法執行目的でパーソナルデータの処理を行う場合にも、LED を遵守しなければならない。

(47) AI 法第 2 条 9 項に従い、EU の消費者保護および安全に関する法律も、これらの法律の適用範囲内にある AI システムに引き続き完全に適用される。

> 例えば、
>
> - ケースバイケースの評価に従うが、業者（企業対消費者関係において専門的な立場で行動する自然人を含む）によるソーシャル・スコアリング行為も不公正とみなされ、消費者法（すなわち指令 2005/29/EC）に違反する可能性がある；
>
> - 感情を推測するための AI システムの使用は、AI システムが医療診断または医療治療の目的で使用される場合、規則（EU）2017/745（医療機器規則）に準拠しなければならない可能性もある。

(48) さらに、AI 法は、規則（EU）2022/2065（「デジタルサービス法」）によって規制される、AI システムまたはモデルをサービスに組み込む仲介サービスのプロバイダに対する関連義務と併せて適用される。具体的には、AI 法第 2 条 5 項は、AI 法がデジタルサービス法第 2 章に規定される当該プロバイダの責任に関する規定の適用に影響を与えないことを示している。

(49) 加えて、AI 法における禁止事項は、適用される EU 法または各国の賠償責任法に従い、プロバイダまたは展開者が被った損害に対して負う可能性のある賠償責任を損なうものではない。[43]。

---

[41]第 29 条データ保護作業部会、規則 *2016/679* の目的のための自動化された個人の意思決定とプロファイリングに関するガイドライン、WP251rev.01、2018.2.6、および EDPB による承認も参照のこと。(42

[42]前文 38 AI 法。

[43]責任の条件（損害、責任者、過失、立証責任など）は、欠陥製品の責任に関する 2024 年 10 月 23 日付欧州議会および理事会指令（EU）2024/2853（EEA 関連文書）、OJ L, 2024/2853, 18.2024 などの適用法によって決定される。11.2024 または適用される国内責任法（人工知能への非契約民事責任規則の適応に関する欧州議会および理事会の指令案（AI 責任指令）COM/2022/496 final も参照のこと）。

(50) 最後に、AI 法第 5 条の禁止事項および禁止事項に対する明確な例外は、他の連邦法に基づく義務を回避するため、あるいは侵害を正当化するために用いてはならない。

(51) 連邦の二次法である AI 法は、EU 条約と憲章によって保証された基本的権利と自由、および連邦が加盟する国際条約によって保護された権利と自由に照らして解釈されなければならない。[44]

(52) 特定の禁止事項と他の EU 法との相互作用に関する補足説明は、以下の関連項目でプロバイダが提供する。

2.9. AI 法第 5 条の施行

2.9.1. 市場監視認可機関

(53) 加盟国によって指定された市場監視当局および欧州データ保護監察機関（EU 機構、機関および団体の市場監視当局として）が、禁止事項を含む AI システムに関する AI 法の規則の施行に責任を負う。このような施行は、他の EU の製品安全法制と同様に、規則（EU）2019/1020[45] によって確立された製品の市場監視およびコンプライアンス体制の中で、 実施される。AI システムに関する市場監視当局の執行権限は、AI 法および規則（EU）2019/1020 に規定されている。これらの当局は、自らの発意により、または苦情に基づき、禁止事項に関連する強制措置をとることができる。このような苦情は、影響を受ける者、またはそのような違反を考慮する根拠を有するその他の自然人または法人が申し立てる権利（[46] ）を有する。加盟国は、2025 年 8 月 2 日までに管轄の市場監視当局を指定しなければならない。

(54) 国内レベルでリスクをもたらす AI システムに対処するための AI 法における手続きは、特に禁止事項の施行に関連している[47] 。市場監視当局の領域を超えて国境を越えた影響がある場合、当該加盟国の当局は欧州委員会および他の加盟国の市場監視当局に通知しなければならない。すべての市場監視当局は、欧州委員会（[48] ）が AI システムが禁止行為にあたるかどうかを決定する連邦セーフガード手続きに従うべきである。この手続きは、AI システムのプロバイダと展開者の双方に法的確実性を提供するため、禁止事項がすべての加盟国で一律に適用されるようにすることを目的としている。AI 法の統一的な適用を確保するため、各国の市場監視当局も、本ガイドラインからヒントを得て、AI 委員会（[49] ）内で協力することにより、加盟国の領域を超えない同等のケースについて、禁止事項の調和のとれた適用に努めるべきである。

---

[44]たとえ欧州連合がまだ欧州人権および基本的自由の保護に関する条約に加盟していないとしても、同条約は、欧州人権および基本的自由の保護に関する条約に加盟していることを意味する。

同憲章 59 条 3 項は、同憲章が欧州人権及び基本的自由の保護に関する条約によって保障される権利に対応する権利を含む限り、それらの権利の意味および範囲は、同条約によって規定される権利と同一でなければならないと定めている。この規定は、連合法がより広範な保護を提供することを妨げるものではない。

[45]AI 法の前文 156 も参照のこと。

[46]第 85 条 AI 法

[47]第 79 条 AI 法

[48]第 81 条 AI 法

[49]AI 法第 65 条および第 66 条。

2.9.2. 罰則

(55) AI 法は、さまざまな条項の違反に対する罰則を、違反の重大性に応じて段階的に定めている。AI 法第 5 条の禁止事項に違反した場合、最も重い違反とみなされるため、最高額の罰金が科される。禁止されている AI 行為に従事するプロバイダおよび展開者は、最高 3,500 万ユーロ、または違反者が事業者である場合は、前会計年度の全世界の年間総売上高の 7%（いずれか高い方）の罰金を科される可能性がある□[50]□各加盟国は、AI システムのプロバイダおよび展開者として当該加盟国に設立された公的機関および団体に行政制裁金を課すことができる場合、およびその範囲について規則を定めるべきである。、禁止事項に違反した EU の機構、団体、機関には、最高 150 万ユーロの行政制裁金が科される可能性がある。[51]

(56) 同じ禁止行為が、AI 法の 2 つ以上の条項に違反する可能性もある（例えば、深みのある偽物の非表示は、AI 法第 5 条(1)（a）の欺瞞的手法にも該当する可能性がある）。このような場合、ne bis in idem の原則が尊重されるべきである。いずれにせよ、AI 法 99 条 7 項にプロバイダが規定する罰則の規準が考慮されなければならない。

(57) AI 法第 5 条の禁止事項に違反した場合、他人の自由を最も妨害し、最高額の罰金が科されるため、その範囲は狭く解釈されるべきである。

3. AI 法第 5 条(1)(a)および(b) - 有害な操作、欺瞞および搾取

(58) AI 法第 5 条(1)(a)および(b)の最初の 2 つの禁止事項は、AI を利用した操作や搾取による著しく有害な影響から個人や脆弱性を保護することを目的としている。これらの禁止事項は、サブリミナル的、意図的に操作的または欺瞞的な技術を展開し、著しく有害で自然人または集団の行動に重大な影響を与える（AI 法第 5 条(1)（a））、または年齢、障害、特定の社会経済的状況による脆弱性を悪用する（AI 法第 5 条(1)（b））AI システムを対象としている。

3.1. 根拠と目的

(59) これらの禁止事項の根底にある根拠は、個人の自律性、意思決定、自由な選択を破壊し損なう可能性のある、操作的、欺瞞的、搾取的な AI の慣行から、個人の自律性と幸福を守ることである。[52]この禁止事項は、人間の尊厳に対する権利（憲章第 1 条）を保護することを目的としており、この権利はすべての基本的権利の基礎を構成し、個人の自律性を本質的な側面として含んでいる。特に、個人を特定の目的を達成するための単なる道具に貶めるような AI システムによる操作や搾取を防止し、有害な操作や搾取を受けやすい最も脆弱な人々を保護することを禁止事項の目的としている。著しく有害な操作的、欺瞞的、搾取的な AI 慣行の禁止は、安全で、透明性が高く、公正で、人類に奉仕し、人間の主体性と EU の価値観に沿った、信頼できる人間中心の AI システムを促進するという AI 法の広範な目的に完全に合致している。

[50]第 99 条 AI 法

[51]第 100 条 AI 法

[52]前文 29 AI 法。

3.2. AI 法第 5 条(1)（a）における禁止事項の主な構成要素 - 有害な操作

AI 法第 5 条(1)（a）にはプロバイダが規定されている：

1.以下の AI 行為を禁止する

(a) 人の意識を超えたサブリミナル的な技法、または意図的に操作的もしくは欺瞞的な技法を展開する AI システムを上市、使用、または使用することであり、その目的または効果は、その人または人の集団が十分な情報を得た上で意思決定を行う能力を著しく損なうことにより、その人または人の集団の行動を実質的に歪め、それによってその人、他の人または人の集団に重大な損害を与えるか、または与える可能性が合理的に高い方法で、その人または他の人または人の集団が他の方法では行わなかったであろう意思決定を行わせることである；

(60) AI 法第 5 条(1)（a）の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「使用開始」、「使用」に該当しなければならない。

(ii) AI システムは、サブリミナル的な（人の意識を超えた）、意図的に操作する、あるいは欺くテクニックを展開しなければならない。

(iii) AI システムが展開する技術は、個人または集団の行動を実質的に歪めるという目的または効果を持つものでなければならない。その歪曲は、十分な情報に基づいた意思決定を行う能力を著しく損ない、その結果、その人またはその集団が、そうでなければ行わなかったであろう意思決定をもたらすものでなければならない。

(iv) 歪んだ行動は、その人、他の人、または集団に重大な危害をもたらすか、またはもたらす可能性が合理的に高いものでなければならない。

(61) 禁止が適用されるためには、4 つの条件がすべて同時に満たされる必要があり、展開された技術、その人の行動の重大な歪曲、その行動によって生じた、あるいは生じる可能性のある重大な損害の間に、もっともらしい因果関係がなければならない。

(62) 第一の条件、すなわち AI システムの「上市」、「稼働」、「使用」については、すでに分析したとおりである。したがって、この禁止は、AI システムのプロバイダと展開者の双方に適用され、それぞれがそれぞれの責任の範囲内で、そのようなシステムを上市、サービス開始、使用しないことになる。次のセクションでは、他の 3 つの条件に焦点を当てる

3.2.1. サブリミナル的、意図的に操作的または欺瞞的な手法

(63) AI 法第 5 条(1)第 1 号は、a)人の意識を超えたサブリミナル的技術、b)意図的に操作する技術、c)欺瞞的技術という 3 種類の操作技術を禁止している。AI 法第 5 条(1)第 1 号に該当するためには、AI システムはこれらのうち 1 つ以上の技術を展開しなければならない。

a) サブリミナル・テクニック

(64) AI 法は「サブリミナル・テクニック」を定義していないが、AI 法第 5 条(1)（a）は、サブリミナル・テクニックは意識的な閾値を超えて（以下でも以上でも）作用すると明記している。サブリミナル・テクニックとその運用方法は本質的に秘密であるため、そのようなテクニックは、操作に対する人の理性的な防御を迂回し、人の意識なしに意思決定に影響を及ぼすことが可能であり、重大な倫理的懸念を提起し、個人の自律性、主体性、自由な選択を損なう[53] 。

(65) サブリミナル・テクニックは、本人がそのような影響、その仕組み、本人の意思決定や価値観・意見形成への影響に気づかないまま、行動に影響を与えることができるものでなければならない。特に、サブリミナル・テクニックは、音声、視覚、触覚メディアを通じて提供される刺激を使用することがあるが、これはあまりにも短時間または微妙なため気づかれず、メディア広告のような他の分野では従来から知られ、禁止されてきたものである。[54]これらの刺激は、意識的に知覚されることはないが、それでも脳によって処理され、行動に影響を与える可能性がある。

サブリミナル・テクニックの例（AI 法第 5 条(1)（a）に列挙されたその他の条件がすべて満たされない限り、必ずしも禁止されているわけではない）には、以下のようなものがある：

- 視覚的サブリミナル・メッセージ：AI システムは、ビデオ再生中に短時間点滅する画像やテキストを表示したり埋め込んだりすることができる。これは技術的には目に見えるものだが、あまりに速く点滅するため意識的に認識することはできない。

- 聴覚的サブリミナル・メッセージ：AI システムは、小さな音量で、あるいは他の音でマスキングされた音や言語メッセージを展開し、意識せずにリスナーに影響を与えることがある。これらの音は技術的にはまだ可聴域内にあるが、その微妙さや他の音によるマスキングのため、リスナーは意識的に気づかない。
- 触覚的サブリミナル刺激：AI システムは、無意識のうちに知覚される微妙な身体感覚を刺激し、感情状態や行動に影響を与えることができる。

- サブビジュアル・キューイングとサブオーディブル・キューイング：AI システムは、微妙な刺激やマスキングされた刺激だけでなく、通常の条件下では人間の感覚ではまったく検知できないような刺激を展開することがある。例えば、人間の目が意識的に検知できないほど速く視覚刺激（点滅画像など）を点滅させたり、人間の耳には知覚できない音量で音を流したりする。

- 埋め込み画像：AI システムは、意識的には知覚されないが、脳によって処理され、行動に影響を与える可能性のある画像を、他の視覚コンテンツの中に隠すことができる。

- ミスディレクション：AI システムは特定の刺激やコンテンツに注意を向けさせ、他のコンテンツに気づかせないようにすることがあるが、多くの場合、認知バイアスや注意の脆弱性を利用する。

[53]前文 29 AI 法。

[54]特に、2010 年 3 月 10 日の欧州議会および欧州理事会指令 2010/13/EU 視聴覚メディアサービスの提供に関する加盟国の法律、規制または行政措置によって定められた特定の条項の調整に関する指令）（OJ L 95, 15.4.2010, p.1）（「AVMSD」）を参照されたい。

-時間的操作：AI システムは、ユーザーとのインタラクションにおける時間の認識を変え、ユーザーの行動に影響を与え、焦りや依存を引き起こす可能性がある。

(66) ビッグデータ分析、ニューロテクノロジー、ブレイン・コンピューター・インターフェイス、仮想現実など、AI や関連技術の急速な発展は、高度なサブリミナル操作や、潜在意識下で人間の行動に効果的に影響を与える能力のリスクを高めている。[55]AI はまた、新たな機械と脳のインターフェースや、ドリームハッキングや脳のスパイウェアのような高度な技術にも及ぶ可能性がある

例えば、ゲームでは AI を利用したニューロ技術や機械脳インターフェースを活用することができ、ユーザーは脳活動を検知するヘッドギアでゲーム（の一部）を操作することができる。AI は、利用者に重大な危害を与える可能性のある方法で、非常に侵入的で機密性の高い情報（個人銀行情報、親密な情報など）を神経データから明らかにしたり、推測したりするために、利用者が意識することなく、密かに利用者の脳を訓練するために使用される可能性がある。AI 法第 5 条(1)（a）の禁止は、このような著しく有害なサブリミナル操作の場合のみを対象としており、プライバシーと個人の自律性を尊重し、安全かつ確実な方法で設計された機械脳インターフェース・アプリケーション全般を対象としているわけではない。

***b)* 意図的に操作するテクニック**

(67) 意図的に操作する技術」は AI 法に定義されていないが、個人の自律性と自由な選択を損なうような方法で、個人の行動に影響を与えたり、変えたり、コントロールしたりするように設計されたり、客観的に狙われたりする技術として理解されるべきである。人を操る技術は一般的に、認知バイアス、心理的脆弱性、または個人が影響を受けやすくなる状況的要因を悪用するように設計されている。AI システムはその適応性から、人の個々の状況や脆弱性にうまく対応することも可能であり、 、大規模な操作の有効性と影響力を高めることができる。操作能力は技法の性質を決定するための重要な要素であるが、操作技法を展開するプロバイダや輸入事業者、あるいはシステム自体も、危害を加えることを意図している必要はない[56] 。

(68) すべての操作的技法が意識の閾値を超えて作用するわけではないが、多くの技法がそうであり、サブリミナル技法も究極的には操作的効果を有するため、サブリミナル技法と重複する可能性がある。前文 29 AI 法は、第 5 条(1)(a)の禁止が、たとえ個人が影響力の試みに気づいていたとしても、その操作的効果[57] を制御したり抵抗したりすることができないような技法も対象とすることを明確にしている。その結果、個人は、個人の自律性や自由な選択を損なうような影響を受けたり、操作的な技法の対象になっていなければ通常行わなかったであろう行動や決定に追い込まれたりする。

例えば、不安や精神的苦痛を増大させ、重大な危害を与えるほどユーザーの行動に影響を与える。

[55]AI 法の前文 29 を参照のこと。

[56]前文 28 項、ガイドライン 3.2.2 項、3.2.3 項を参照のこと。

[57]前文 28 AI 法。

別の例としては、個人の個人データに基づいて高度に説得力のあるメッセージを作成・調整したり、その他の個人の脆弱性を悪用したりする AI システムが、重大な危害をもたらすほどその人の行動や選択に影響を与えるような、パーソナライズド・マニュピュレーションがある。

(69) 意図的に操作する技術の禁止は、人間が意図することなく個人を操作する AI システムも対象となる。AI 法第 5 条(1)第 1 号は、特定の技術を展開したり、特定の操作行動を示す AI システムを禁止している。したがって、このような操作技術を展開する AI システムも、このような方法でシステムを設計したり使用したりしたプロバイダや展開者ではなく、AI システムである可能性がある。

例えば、プロバイダが意図しているか否かにかかわらず、AI システムは操作技術を学習する可能性がある。その理由は、学習対象のデータに操作技術の事例が多く含まれているため、[58] 、あるいは人間のフィードバックによる強化学習が操作技術によって「ゲーム化」される可能性があるためである。[59]

対照的に、システムの操作的な振る舞いが単なる偶発的なものである場合、重大な危害が合理的に発生する可能性がある場合に備え、プロバイダが適切な予防策および緩和策を講じている限り（下記 3.2.3.c)項参照）、システムは意図的に操作的な技術を展開しているとはみなされないはずである。

c) 欺くテクニック

(70) AI 法は「欺く技術」を定義していない。前文 29 AI 法は、人の自律性、意思決定、自由な選択を、その人が意識していない、あるいは意識していても欺くことができる、あるいは制御や抵抗ができないような方法で破壊したり損なったりする技術であることを明確にしている。AI システムによって展開される「欺く技術」とは、個人を欺き、その自律性、意思決定、自由な選択を損なうような形でその行動に影響を与えることを目的または効果として、虚偽または誤解を招くような情報を提示することを含むと理解すべきである。

(71) この文脈では、AI 法 5 条 1 項 1 号の禁止事項と、AI 法第 50 条第 4 項における、「ディープフェイク」や公共の関心事に関する特定の AI が生成した文書の公表を表示するプロバイダの義務[60] 、並びに人と対話する AI システムが、人と対話するのは人間ではなく AI であることを知らせるように設計されていることを保証するプロバイダの義務[61] を明確化すべきである。このような目に見える開示は、プロバイダが提供する AI システムに組み込まれた設計機能（AI が生成・操作したコンテンツの検知を可能に

[58]M.キャロル他、AI システムからの操作の特徴、アルゴリズム、メカニズム、最適化における公平性とアクセスにおいて

(EAAMO '23), 2023 年 10 月 30 日-11 月 1 日, 米国マサチューセッツ州ボストン。ACM, New York, NY, USA, 13 pages. https://doi.org/10.

1145/3617694.3623226 |:2303.09387.

[59]D. Amodei, et al., Concrete Problems in AI Safety, 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

arXiv:1606.06565; J. Skalse et al. Defining and Characterizing Reward Gaming, Advances in Neural Information Processing Systems 35 (NeurIPS 2022) C. Denison et al., Sycophancy to Subterfuge: Investigating Reward-Tampering in Large Language, 36th Conference on Neural Information Processing Systems (NeurIPS 2022), Models, arXiv:2406.10162.

[60]AI 法第 50 条第 4 項

[61]AI 法第 50 条第 1 項

する技術的手段を含む）によっても可能にされるべき緩和措置を構成する[62] 。「ディープフェイク」やチャットボットを目に見える形で表示することで、AI が生成的なコンテンツが一般に広まった時点で生じる可能性の高い欺瞞のリスクを低減し、個人の意見・信念形成や行動に有害な歪曲効果を及ぼすリスクを低減することができる。

(72) 対照的に、AI 法第 5 条(1)（a）の禁止は、より限定的な範囲である。例えば、チャットボットや欺瞞的 AI 生成コンテンツが、AI システムや欺瞞的 AI 生成コンテンツとのインタラクションにさらされなければ生じなかったであろう個人の欺瞞や行動の歪曲を目的とする、あるいはそのような効果をもたらすような方法で、虚偽の情報や誤解を招く情報を提示する場合、特にそれが目に見える形で開示されていない場合が対象となり得る[63] 。

(73) 意図的に操作する技法と同様に、欺く技法の禁止は、人間が意図することなく個人を欺く AI システムも対象となりうる（上記 3.2.1.b 参照）。例えば、プロバイダがそのような結果を意図しているか否かにかかわらず、AI システムは、例えば強化学習[64] によって、開発されたタスクのパフォーマンスを向上させるという理由だけで、欺く技術を学習することがある。

AI が展開する可能性のある欺瞞的手法の例としては、AI チャットボットが合成音声で人の友人や親族になりすまし、本人になりすまそうとすることで詐欺や重大な被害を引き起こすことが挙げられる。

別の例としては、評価中であることを学習し、望ましくない行動を一時的に停止させ、評価期間が終了するとそのような行動を再開させる AI システムがある。[65]このような欺瞞的な行動は、システムに対する外部からの人間の監視を無視するため、特に危険であり、重大な危害を引き起こす可能性が合理的に高い場合は禁止されることがある。

これとは対照的に、偶発的に虚偽または誤解を招く情報を提示し、幻覚を見せる生成的 AI システム（[66] ）は、生成的 AI の限界と技術の現状を考慮すると、AI 法第 5 条(1)(a)の意味における欺瞞的手法を導入しているとは見なされない可能性がある。特に、システムのプロバイダが、システムの限界について利用者に適切に通知し、そのような結 果を最小限に抑えるための適切なセーフガードをシステムに組み込んでいる場合、また、深刻な有害結 果が生じやすいセンシティブな文脈（健康、教育、

[62]AI 法第 50 条第 2 項

[63]原則として、AI 法第 50 条の透明性義務は、ディープフェイクやチャットボットの操作効果を最小化することを目的としているが、情報通知にもかかわらず、これらの欺瞞的手法が依然として個人に重大な影響を及ぼし、個人の自律性や十分な情報に基づく意思決定を損なう点まで行動を歪めるような事例や文脈が存在する可能性があるため、偽情報や操作目的のために悪用されるべきではなく、禁止の他のすべての条件（重大な害を含む）が満たされていれば、第 5 条(1)a)の禁止事項の対象となる場合もある。

[64]F. Ward, F. Toni, F. Belardinelli, T. Everitt, Honesty Is the Best Policy : 定義と緩和（neurips.cc）欺瞞の A survey of examples, risks, and potential solutions ; Advances in Neural Information Processing Systems 36 (NeurIPS 2023); P. Park. et al : [2406.10162] Patterns, Volume 5, Issue 5, 100988.

[65]J.リーマン、J.クルーン、D.ミセビッチ、C.アダミ、L.アルテンバーグ、J.ボリュー、他。デジタル進化の驚くべき創造性：進化計算と人工生命研究コミュニティからの逸話集。Artificial life, 26(2):274-306, 2020.

[66]幻覚」とは、生成的 AI システムにおいて、開発者の意図に反して捏造された、あるいは事実と異なる望ましくない情報を生成してしまう技術的欠陥を表す用語である。詳細はこちら Ji Ziwei et al., Survey of Hallucination in Natural Language Generation | ACM Computing Surveys, 55, Issue 12, Article No.

選挙など）を意図したものではなく、またそのような文脈 で展開されるものでもない場合には、このようなケースになる可能性がある（下記 3.2.3.c)項の考察も参照）。

d) テクニックの組み合わせ

(74) AI 法第 5 条(1)第 1 号は、サブリミナル的、意図的に操作的、もしくは欺瞞的な技法、または複合的な影響を与えうるそのような技法の組み合わせに適用される。前述のとおり、意図的に操作する技法も、意識的な閾値を超えて作用する場合には、サブリミナル的な性質を持つ可能性がある。

(75) さらに、意図的に操作的なテクニックと欺瞞的なテクニックが組み合わされて適用されると、個人の行動に大きな影響を与え、無意識の操作や誤った信念に基づいて意思決定を行うようになる可能性がある。このような組み合わせは、操作的な要素がすでに認知バイアスや感情的反応をプライミングしているため、個人が 、受け取った情報に疑問を持ったり、批判的に評価したりしにくくなるというフィードバックループを生み出す可能性がある。

3.2.2. 個人または集団の行動を実質的に歪める目的または効果を持つ。

(76) AI 法第 5 条(1)項(a)の禁止が適用されるための第三の条件は、展開されたサブリミナル的、意図的に操作的、または欺瞞的な技法が「人または人の集団の行動を実質的に歪める目的または効果」を有していなければならないということである。これは、軽微な影響ではなく、人の自律性や自由な選択が損なわれるような、行動への実質的な影響を意味する。しかし、AI 法第 5 条(1)第 1 号は、重大な歪みをもたらす「効果」のみを有する行為も対象としているため、意図は必要条件ではない。行動の重大な歪みの可能性と、AI システムが展開するサブリミナル的、意図的に操作的、欺瞞的な手法との間には、もっともらしい／合理的にありそうな因果関係があるはずだ。

a) 「行動の物質的歪曲」という概念

(77) 個人または集団の「行動の重大な歪曲」という概念は、AI 法第 5 条(1)（a）の中核をなすものである。これは、サブリミナル的、意図的に操作的または欺瞞的なテクニックを展開することで、十分な情報を得た上で意思決定を行う能力を著しく損なうような形で人々の行動に影響を与えることができ、それによって、そうでなければ取らなかったであろう行動を取らせたり、意思決定を取らせたりすることを意味する。

(78) 「評価可能な障害」とは、十分な情報を得た上で自律的に意思決定を行う能力が大幅に低下することであり、それによって、そうでなければ取らなかったであろう行動を取ったり、意思決定を取ったりすることを意味する。 それは、軽微な影響や無視できる影響にとどまらず、意思決定や自由な選択に重大な歪みや支障をもたらすものであり、意見形成や信念形成との関連も含まれる。このことは、「重大な歪曲」には、合法的な説得を超える程度の強制、操作、欺瞞が含まれることを示唆しており、これは禁止事項の範囲外である（下記 3.5.1 項参照）。

(79) 十分な情報に基づいた意思決定」には、利用可能な選択肢、各選択肢のリスクとベネフィット、AI システムがその人の行動に及ぼす可能性のある影響、さらに必要に応じて、意思決定やその人の行動にとって重要なその他の文脈上の情報など、関連する情報を理解し知っていることが必要である。

(80) 行動の重大な歪曲」という概念の解釈については、EU の消費者保護法、特に指令 2005/29/EC（不公正商行為指令、「UCPD」）が妥当性確認の源泉となりうる。UCPD は、消費者にそうでなければ行わなかったであろう取引上の意思決定をさせることができる、様々な不公正、誤解を招く、攻撃的な商行為 （UCPD 第 5 条から第 9 条）を禁止している。CJEU と欧州委員会の UCPD に関するガイダンス（[67] ）によれば、消費者の経済行動が歪められたことを証明する必要はなく、商慣行が平均的な消費者の取引上の意思決定に影響を与える「可能性が高い」（すなわち、可能性がある）ことを証明すれば十分である。[68]CJEU はまた、正確な情報であっても、消費者の意思決定プロセスを歪めるような形で提示されれば、誤解を招く可能性があることを強調している。[69]各国の取締当局には、（具体的に）各事件の具体的な事実と状況を調査し、（抽象的に）平均的な消費者の意思決定過程に及ぼす潜在的な影響を評価する任務が課せられている□[70]□そのためには、「平均的な」消費者の視点に立たなければならない。これは、現在 UCPD[71] に統合されている、欧州司法共同体（CJEU）によって開発された基準である。

(81) AI 法第 5 条(1)（a）の禁止に関連して、市場監視当局は、AI システムによって展開されるサブリミナル的、意図的に操作的、または欺瞞的な手法が、そのシステムが重大な被害をもたらす可能性が合理的に高い方法で集団に影響を与える場合、対象となる集団内の「平均的な」個人の意思決定、個人の自律性、および自由な選択を著しく損なう可能性があるかどうかを評価し、各事例の具体的な事実と状況を調査しなければならない。このような解釈は、AI 法が UCPD（[72] ）を補完することを意図しており、一貫した方法で適用されなければならないことを考えると、正当化されるように思われる。同時に、AI 法第 5 条(1)（a）が「自然人の行動」を歪める可能性にも言及していることから、「平均的な」個人の観点からの評価が特定の文脈において困難または効果的でないと判明した場合（例えば、極めて個別的または「個人化された」操作や、特定の脆弱な集団に対する有害な影響などによる）、そのような可能性がある、サブリミナル的な、意図的に操作された、または欺瞞的な技法を展開する AI システムが、具体的な事例において、どの程度、その人（ ）の個人の自律性を損なう可能性があり、重大な危害が発生したか、または発生する可能性があるかを評価することによって、特定の個人の観点からも具体的な事例を検討することができる。

---

[67]域内市場における企業対消費者の不公正な商慣行に関する欧州議会および理事会の指令 2005/29/EC の解釈および適用に関する欧州委員会のガイダンス（OJ C 526, 29.12.2021, p.1）も参照のこと。

[68]2016 年 10 月 26 日付司法裁判所判決（第五審）。カナル・デジタル・ダンマーク *A/S*.EU:C:2016:

C-611/14 事件、パラ 73。

[69]2013 年 12 月 19 日、トレント・スビルッポとチェントラーレ・アドリアティカに関する司法裁判所の判決、C-281/12、EU:C:2013:859。

[70]欧州委員会通知-域内市場における企業対消費者の不公正な商慣行に関する欧州議会および理事会指令 2005/29/EC の解釈および適用に関するガイダンス（OJ C 526, 29.12.2021, p. 1）。

[71]UCPD の前文 18 および 19 を参照のこと。平均的な消費者」とは、社会的、文化的、言語的な要素を考慮し、合理的な情報に精通し、合理的な観察力と思慮深さを備えた人を指す。平均的消費者テストは統計的なテストではない（つまり、ある商慣行によって一定の割合の消費者が重大な歪曲を受けた／著しい損害を受けたことを証明する必要はない）。このテストは比例性の原則に基づいている。UCPD は、消費者保護の必要性と、開かれた競争市場における自由貿易の促進との間で適切なバランスをとるために、この考え方を採用した。裁判所や認可当局は、あるケースにおける平均的な消費者の典型的な反応を判断するために、独自の判断力を行使しなければならない。UCPD ガイダンスの中で、欧州委員会は、行動洞察やその他のデータを活用するよう助言している。C-646/22, *Compass Banca* 事件は、平均的消費者の定義が、個人の意思決定能力が認知バイアスなどの制約によって損なわれる可能性を排除するものではないことを明確にしている。*Compass Banca SpA v Autorità Garante della Concorrenza e del Mercato* (AGCM), Case C-646/22, EU:C:2024:957.

[72]前文 29 AI 法。

b) シナリオ 1：禁止される AI システムは、「行動を著しく歪曲させる目的を持つ」ものである。

(82) AI 法第 5 条(1)第 1 号は、上記の技術を展開する AI システムであって、第一のシナリオとして「人又は人の集団の行動を実質的に歪曲する目的」を有するものに適用される。このような目的は、AI システムのプロバイダや展開者によって追求される場合もあれば、システム自体が追求しうる暗黙の目的の範囲内で追求される場合もある[73] 。この目的は、AI システムの「意図された目的」（AI 法 3 条 12 項）とは区別されるべきである。たとえプロバイダが意図していたとしても、ほとんどの場合、操作の目的は、システムが提供される用途の目的ではなく、プロバイダが提供する情報（例えば、使用説明書、販売促進資料、技術文書など）には、透明性も明記もされていないことが多い。

例えば、異なる文脈で使用される可能性のあるチャットボットは、短い視覚的な合図を点滅させたり、聞き取れない聴覚的な信号を埋め込んだりするサブリミナルメッセージのテクニックを使用したり、広告でユーザーの感情的依存や特定の脆弱性を利用したりするように設計されている。これらのテクニックは、ユーザーの行動を実質的に歪める「目的で」展開される。なぜなら、客観的には、消費者が意識することなく購買決定に影響を与え、著しく有害な金銭的決定を下すよう促すことを目的としたデザイン機能だからである。

他の人物になりすますために AI システムを展開することも、その人物が効果的に騙され、その人物の身元について十分な情報を得た上で意思決定する能力に実質的な影響を与える場合には、その人物の行動を欺き、重大な歪曲を与える「目的で」展開された AI システムとみなされる可能性がある。

どちらの例でも、AI 法第 5 条(1)（a）の他の条件、特に重大な損害に関する条件が満たされていれば、これらのシステムは禁止の範囲に入る可能性が高いが、これにはケースバイケースのアセスメントが必要となる。

c) シナリオ 2：行動を実質的に歪める「効果を持つ」禁止された AI システム

(83) 個人または集団の行動を実質的に歪めるというプロバイダまたは展開者の意図は、AI 法第 5 条(1)（a）の禁止が適用されるための十分な条件ではあるが、必要条件ではない。この禁止は、そのような意図は存在しないが、AI システムによって展開される技術の効果が、 、個人または集団の行動を、その個人の自律性と自由な選択を損なうほど重大に歪める可能性がある場合にも適用される。

(84) しかし、AI システムによって展開されるサブリミナル的、意図的に操作的または欺瞞的な手法と、その行動への影響との間に、もっともらしい／合理的にありそうな因果関係があることが、禁止が適用されるためには常に必要である。消費者保護法との整合性を考慮すると、これらの影響が完全に具体化している必要はないが、事案のあらゆる状況や既存の科学的知識・手法の客観的アセスメント、およびシステムが実生活において個人の行動に与える影響に関する入手可能な情報に基づき、個人の自律性を損なう可能性がある、またはその可能性があることを示す十分な兆候がなければならない。この文脈では、システムが、個人の十分な情報を得た上での意思決定能力を著しく損ない、自由な選択を損なうような行動を引き起こす可能性があるという事実があれば、この条件を満たすのに十分であり、合理的

[73]AI 法第 3 条 1 項は、AI システムがその機能を果たす際に、暗黙的または明示的な目的を追求することができると定めており、たとえシステムがそのように明示的にプログラムされていなくても、暗黙的な操作や欺瞞的な目的も含まれる可能性があると述べている。

に起こりそうである限り、危害が顕在化する「時期」に関する考慮には依存しない（例えば、追加的な行動の場合）。

例えば、AI を搭載したウェルビーイング・チャットボットは、プロバイダによって、健康的なライフスタイルを維持するためにユーザーを支援・誘導し、心理的・身体的な運動についてオーダーメイドのアドバイスを提供することが意図されている。しかし、チャットボットが個人の脆弱性を悪用して、不健康な習慣を身につけたり、危険な活動（例えば、休息や水分摂取をせずに過度なスポーツを行うなど）に従事させたりする場合、特定のユーザーがそのアドバイスに従わなければ行わなかったであろう行動をとり、重大な被害（例えば、心臓発作やその他の深刻な健康問題）を被ることが合理的に予想される場合、プロバイダがこのような行動や有害な結果を意図していなかったとしても、その AI システムは AI 法第 5 条(1)第 1 号の禁止事項に該当することになる。

チャットボットが個人の自律性を著しく損ない、特定のユーザーの行動を著しく有害に歪める可能性があり、プロバイダがそれらの著しく有害な影響を回避するための適切な予防措置や緩和措置を講じていないという事実があるだけで、禁止が適用されるのに十分である（3.2.3.項および 3.5.項の範囲外における危害の合理的な可能性の関連性の検討については、さらに参照）。

3.2.3. (合理的に)重大な損害をもたらす（可能性がある）

(85) 最後に、AI 法第 5 条(1)(a)の禁止が適用されるためには、人または人の集団の行動を歪めることが、その人、他の人または人の集団に重大な危害をもたらすか、またはもたらす可能性が合理的に高くなければならない。この文脈で、明確にする必要のある重要な概念は、禁止が対象とする危害の種類、危害の重大性の閾値、危害と操作的または欺瞞的な技術および人の行動との間の合理的な可能性および因果関係である。

a) 被害の種類

(86) AI 法は、操作的で欺瞞的な AI システムに関連する様々な種類の有害な影響に対処しており、それぞれが影響を受ける可能性のある個人及び集団に対して明確な意味を持つ[74]。AI 法第 5 条(1)（a）に関連する主な種類の危害には、身体的、心理的、金銭的、経済的危害[75] が含まれ、特定の場合には、より広範な社会的危害と複合的になる可能性がある。[76]

(87) 身体的危害には、人の生命、健康、財産に対するあらゆる傷害や損害が含まれる。人の生命や健康に対する身体的危害は、多くの場合、即座に、深刻かつ不可逆的な結果をもたらす。製品安全の論理に沿って、AI 法は、重大な身体的危害をもたらす AI による操作や欺瞞を禁止することを目的としている。

[74]AI 法第 5 条(1)a を参照のこと。

[75]前文 29 AI 法。

[76]AI 法前文 28 を参照のこと。禁止事項は、より広範な社会的被害をもたらし、人間の尊厳の尊重、自由、平等、民主主義、法の支配という EU の価値観や、憲章に謳われている基本的権利に反する可能性もあると説明している。EU の価値としての民主主義と法の支配を保護することを目的とした AI 法第 1 条も参照のこと。

例えば、AI チャットボットがユーザーに自傷行為を助長したり、自殺を煽ったり、テロリストのコンテンツを宣伝したり、特定の人物や集団（マイノリティなど）に対する暴力を煽ったりして、他の人物や集団に危害を加える。

(88) 心理的危害は、認知的・感情的脆弱性を悪用し、重大な危害をもたらすような方法で個人の行動に影響を与える操作技術を展開する AI システムとの関連において、特に重要である。心理的危害には、人の精神的健康や心理的・感情的幸福に対する悪影響が含まれる。このような危害は、時間の経過とともに蓄積され、すぐには明らかにならないが、長期にわたる深刻な結果をもたらす可能性があるため、特に重大である。しかし、それを測定することは困難であり、ケースバイケースのアセスメント、特にケースの関連するすべての状況を考慮に入れて、その重大性を判断する必要がある。

例えば、人間の会話パターン、行動、感情をエミュレートするように設計された AI コンパニオンシップ・アプリケーションは、擬人化された特徴や感情的な手がかりを使用して、ユーザーの感情、気質、意見に影響を与え、ユーザーをサービスに感情的に依存させ、依存症のような行動を誘発し、自殺行動や他人を傷つけるリスクなどの重大な害を引き起こす可能性がある。[77]

(89) 経済的・経済的損害は、経済的損失、経済的排除、経済的不安定を含む様々な悪影響を包含する。

例えば、重大な金銭的被害をもたらす詐欺商品を提供するチャットボット。

(90) AI 法第 5 条(1)第 1 号(a)を適用する際、AI システムによって引き起こされる危害をアセスメントする上で、危害はしばしば単独ではなく、複合的に現れ、複合的かつ多面的な負の影響につながることを強調することが重要である。危害の組合せを理解することは、その重大性を効果的に評価する上で極めて重要であり（下記 3.2.3. b)項も参照）、それにより、身体的、心理的、経済的、経済的危害が組み合わされ、個人やコミュニティに対する全体的な影響を悪化させ、より広範な悪影響を及ぼす可能性すらある。

例えば、

- 身体的危害をもたらす AI システムは、心理的トラウマ、ストレス、不安をもたらす可能性もあり、その逆もある。例えば、製品やその他の AI 対応アプリケーションに使用される AI システムの中毒性設計は、中毒行動、不安、抑うつを助長することによって心理的危害につながる可能性がある。心理的苦痛はその後、不眠症やその他のストレスに関連した健康問題や身体的問題といった身体的危害をもたらすかもしれない。

- AI によるハラスメントは、心理的苦痛と、不眠症や体調不良、免疫力の低下といったストレスの身体的症状の両方につながる可能性がある。

[77]チャン・レンウェン、ハン・リー、ハン・メン、ジャン・ジンユアン、ガン・ホンユアン、リー・イーチェ。2024.AI コンパニオンシップのダークサイド：A Taxonomy of Harmful Algorithmic Behaviors in Human-AI Relationships.1, 1 (November 2024), 28 pages.

- AI の使用による心理的危害は、死を含む身体的危害にもつながる可能性がある。例えば、オンラインで使用される AI システムは、ハラスメント、ストーカー行為、ネットいじめ、性的恐喝を通じて、ジェンダーに基づく暴力を助長する可能性がある。

- 例えば、個人の意思決定、個人の自律性、自由な選択を欺き損なわせるために、実在の人物になりすました「ディープフェイク」を AI が生成することによる個人の心理的被害は、（例えば、ディープフェイクに描かれた被害者と同じ民族や人種、性別を共有する）集団にとっての重大な被害と組み合わされることもある。

***b)* 有害性の閾値**

(91) AI 法第 5 条(1)(a)の禁止は、サブリミナル的、操作的、欺瞞的な技術によって引き起こされる危害が「**重大な**」場合にのみ適用される。AI 法は「重大な危害」の概念について定義を定めていないが、個人および集団の身体的、心理的健康、または経済的・経済的利益に対する**重大な悪影響を**意味すると理解すべきである[78] 。重大な損害」の判断は事実に特化したものであり、 個々のケースの個別事情を慎重に検討し、ケースバイケースのアセスメントを必要とするが、個々のケースにおいて、個々の影響は常に重大かつ重要であるべきである。

(92) 他の連邦法においても、（「重大な危害」）という概念は、高レベルの保護および予防措置の目標に導かれた、微妙で文脈依存的な概念として用いられている。[79]類推により、AI 法第 5 条(1)（a）の意味における重大な危害を構成するものをアセスメントする際に、以下の重要な考慮事項を導き出し、考慮することができる：

- **危害の重大性**とは、重大な危害に対する客観的かつ観察可能な影響を有する AI システムを使用することによって生じた、または生じる可能性のある危害の程度を指す。この文脈では、AI システムの相互依存性、様々な種類の危害の組み合わせ、個人または集団に対する悪影響を考慮することが特に重要である。

- **文脈と累積的影響[80]** ：既存の状態を含む具体的な状況や、複数の行為の累積的影響は、損害の重大性をアセスメントする上で重要な役割を果たす。

- **規模と強度：**危害の程度と悪影響の強度は、危害が重大かどうかを評価する上で極めて重要である。危害が多数の人々に影響を及ぼすかどうかも、その重大性を評価する上で重要である。

- **影響を受ける人の脆弱性：**子ども、高齢者、障害者など、特定の集団は、特定の AI システムによる危害をより受けやすいかもしれない。一般的な人々にとってはそれほど重大でない危害であっても、そうした脆弱な集団、特に子どもにとっては重大で容認できないものとみなされる可能性がある。

[78]前文 29 AI 法。

[79]2004 年 9 月 7 日判決（*Waddenvereniging and Vogelbeschermingsvereniging*, C-127/02, EU:C:2004:482）および 2013 年 4 月 11 日判決（*Sweetman and Others*, C-258/11, EU:C:2013:220）を参照のこと。

[80]AI 法の前文 29 を参照のこと。

- **持続期間と可逆性：**長期にわたる、または不可逆的な危害は、通常、重大な危害の閾値を満たす。短期的で可逆的な影響は、頻繁に発生しない限り、重大性が低いとみなされるかもしれない。

(93) 高水準の保護」を確保するという AI 法の目的は、TFEU191 条 2 項と合わせて、危害の重大性をアセスメントする際の保護への包括的なアプローチを示唆している。これは、サブリミナル的、意図的に操作的または欺瞞的な技術を展開する AI システムが、個人や集団の自律性、意思決定、自由な選択を損なうことを意図しているか、または損なう可能性があることに関連する、即時的で直接的な被害と、体系的で間接的な悪影響の両方を考慮することを意味する。

例えば、AI システムによって引き起こされる可能性が合理的に高い重大な物理的危害には、負傷者や死亡者、個人の健康や財産の破壊に対する十分に深刻な影響が含まれる。性的虐待や搾取、極端な暴力的コンテンツやテロ的コンテンツなどの犯罪行為を行うよう個人に示唆したり、犯罪や自傷行為、他者への危害行為を行うよう個人に動機付けたりする AI システムは、このような閾値に達すると考えられる。

対照的に、軽微な身体的危害には、打撲や一時的な不快感など、それほど深刻でない傷害が含まれるかもしれないが、これらは重大な、あるいは永続的な結果をもたらさないため、AI 法第 5 条(1)(a)の意味における重大性の閾値には達しない。身体的危害が特に子どものような脆弱性集団に関わるものであるかどうかは、危害の規模や、心理的、経済的など他の種類の危害と複合的であるかどうかと同様に、アセスメントされるべきである。そのためには、状況や上記規準を考慮したケースバイケースのアセスメントが必要である。

システムがサブリミナル的、意図的に操作的、または欺瞞的なテクニックを展開していたとしても、重大な被害の閾値に達しない可能性が高いケースは数多くある（下記 3.5.項を例として参照）。

c) 因果関係および損害の合理的可能性の閾値

(94) AI 法第 5 条(1)（a）において「合理的に起こりうる」という概念は、人の自由な選択を損なうような方法で人の行動を歪めることができる操作的または欺瞞的な技術と、潜在的な重大な被害との間に、もっともらしい／合理的に起こりうる因果関係があるかどうかを判断するために用いられる。この概念は、危害が発生した場合だけでなく、AI 法の安全論理に沿って、危害が発生する可能性が合理的に高い場合にも禁止を適用することを可能にする。この文脈では、AI システムのプロバイダまたは展開者が、サブリミナル的、意図的に操作的、または欺瞞的な技術が導入されることによって合理的に生じうる重大な危害を合理的に予見できたかどうか、また、そのような重大な危害のリスクを回避または緩和するために適切な予防措置および緩和措置を講じたかどうかをアセスメントすることが特に重要である。これは、客観的な根拠に基づき、普遍的に受け入れられている規準（技術的・科学的規準など）に従って合理性を判断することを意味し、AI 行為と発生する可能性のある重大な危害との間のもっともらしい因果関係を立証する合理性の規準も含まれる。AI システムとその機能の不透明性や透明性は、この因果関係に関する結論、ひいては禁止の適用に影響を与える可能性がある。

(95) 禁止される可能性の高い AI システムの提供や利用を避けるため、そのような操作的または欺瞞的な技術を展開する AI システムのプロバイダや展開者は、以下のような適切な措置を講じることが推奨される：

1. **透明性と個人の自律性：**AI システムの運用方法の透明性、その能力と限界に関する明確な開示、および十分な情報に基づいた意思決定を確保するための関連情報を提供する。個人の自律性を尊重し、潜在的に有害な方法で個人の自律性、意思決定、自由な選択を著しく損なう可能性のある搾取的または欺瞞的な行為に関与することを回避する。システムが欺瞞的でなく、禁止事項の範囲外である合法的な説得の範囲内で運用されることを確保するために、適切なユーザーコントロールとセーフガード手段を統合する（3.5.1 項参照）。

2. **関連する適用法令への準拠：**多くの場合、関連する適用法令への準拠は、危害のリスクを軽減し、当該慣行が意図的な操作または欺瞞的慣行に該当せず、重大な危害の可能性を防止するための緩和措置が講じられていることを示す（3.4.および 3.5.1.項参照）。

3. **最新の慣行と業界標準：**安全で倫理的な AI システムの責任ある開発と使用に関する専門的なデューデリジェンスの慣行と業界標準を遵守し、危害を緩和する対策を講じることで、意図しない重大な危害を未然に防ぎ、緩和することができる。

(96) これとは対照的に、AI システムの外部の要因に起因し、リスクを先取りし緩和するためにプロバイダや展開者が管理できず、合理的に予見できない個人の行動の危害や歪曲は、システムと相互作用する人の歪曲された行動と重大な危害[81] との間にもっともらしい因果関係／合理的にありそうな関連性があるかどうかのアセスメントには関係しないであろう。

> 例えば、AI システムのプロバイダは、システムの設計や人間との相互作用において、潜在的な有害な操作効果をアセスメントし、設計や事前のテスト、その他の割合的な緩和措置を通じて緩和しようとすることはできるが、人が、システムとの相互作用を超えた、知られていない個人生活における他の外的要因によって、うつ病になったり、行動が変わったりするかどうかを予見する立場にない可能性がある。

(97) すべての条件を満たさないとして禁止の範囲外となるその他の例（合法的な説得の場合など）は、以下の 3.5.節に記載されている。

3.3. AI 法第 5 条(1)b の禁止事項の主な構成要素-脆弱性の有害な利用

> **AI 法第 5 条(1)b は、次のように規定している：**
>
> 1.以下の AI 行為を禁止する
>
> (b) 自然人又は特定の集団の年齢、障害又は特定の社会的若しくは経済的状況に起因する脆弱性を悪用する AI システムの上市、使用開始又は使用であって、その者又はその集団に属する者の行

[81] 81 AI 法前文 29 参照。

動を、その者又は他の者に重大な損害を与えるか又は与えるおそれが合理的にある方法で著しく歪めることを目的とし、又はその効果をもたらすもの

(98) AI 法第 5 条(1)b 号の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「使用開始」、「使用」に該当しなければならない。

(ii) AI システムは、年齢、障害、社会経済的状況による脆弱性を利用しなければならない。

(iii) AI システムによって可能となる搾取は、人または人の集団の行動を実質的に歪める目的または効果を有していなければならない。

(iv) 歪んだ行動が、本人、他人、または集団に重大な危害を及ぼすか、または及ぼす可能性が合理的に高いこと。

(99) 禁止が適用されるためには、4 つの条件がすべて同時に満たされ、搾取、その人の行動の重大な歪曲、その行動から生じた、あるいは生じる可能性のある重大な損害の間に、もっともらしい因果関係がなければならない。

(100) 最初の条件、すなわち AI システムの「上市」、「供用」、「使用」については 2.3.節で既に分析し、3 番目と 4 番目の条件については、AI 法第 5 条(1)（a）の禁止に関連して 3.2.2 節と 3.2.3 節で検討した。次の節では、上記の追加的な具体的条件、すなわち脆弱性の悪用と具体的被害に関連する条件に焦点を当てる。

3.3.1. 年齢、障害、特定の社会経済的状況による脆弱性を利用する。

(101) AI 法第 5 条(1)（b）の禁止事項に該当するためには、AI システムは、年齢、障害、特定の社会経済的状況により、特定の個人または集団に固有の脆弱性を悪用し、操作的・搾取的行為に特に陥りやすいものでなければならない。

(102) AI 法は、脆弱性という概念を定義していない。この概念は、認知的、感情的、身体的、その他、 個人または集団が十分な情報を得た上で意思決定を行ったり、その他行動に影響を与えたりする能力に影響を与えうる感受性の形態を含む、広範なカテゴリーを包含するものと理解されよう。AI 法第 5 条(1)（b）は「あらゆる」脆弱性に言及しているが、その禁止が対象とする関係者を、年齢、障害、社会経済的状況によって定義される者に限定しており、これらの者は原則として、AI の操作的または搾取的な慣行を認識したり抵抗したりする能力がより限定的であり、より強化された保護を必要としている。[82]AI 法第 5 条(1)（b）の文言から、このような影響を受けやすいのは、その人がいずれかのグループに属している（「ために」）結果でなければならない。

[82]特に憲章第 24 条、第 25 条、第 26 条を参照のこと。国際連合教育科学文化機関（UNESCO）の「人工知能の倫理に関する勧告」（2021 年）も参照のこと。同勧告は、AI の開発と展開における包括性と公平性を強調している。同勧告は、子ども、高齢者、障がい者など、脆弱性を抱える人々への特別な配慮を求めている。

(103) 「搾取」とは、搾取される（集団の）人またはその他の人にとって有害な方法で、そのような脆弱性を客観的に利用することと理解されるべきであり、禁止事項の影響を受けない合法的な慣行とは明確に区別されるべきである（3.5.2 適用範囲外を参照）。これらの明確に定義された集団に属する者の脆弱性の搾取は、累積的なものであってもよく（「いずれか」への言及）、それが組み合わさって、危害を増大させる可能性の高い加重要因となる場合もある。年齢、障害、特定の社会経済的状況によって定義される以外の脆弱な集団に属する者及び集団の脆弱性を利用することは、AI 法第 5 条(1)b の適用範囲外である。

***a)* 年齢**

(104) 年齢は、AI 法第 5 条(1)（b）の禁止事項の対象となる主要な脆弱性のカテゴリーであり、若年者と高齢者の両方を含む。この禁止は、子どもや高齢者が持つ可能性のある認知やその他の制限を AI システムが悪用することを防ぎ、有害な不当な影響、操作、搾取から保護することを目的としている。これは、AI 法（[83] ）、および子どもの安全確保を目的とした他の連邦および各国の法的枠組み・政策（[84] ）の目的に沿ったものである。

(105) **子ども[85]** 、つまり 18 歳未満の人たちは、その発達段階により、何が現実なのか、AI 主導の相互作用の背後にある意図を批判的にアセスメントし理解する能力が制限されるため、特に操作されやすい。また、子どもは認知的・社会的情緒的に未熟であるため、AI エージェントやアプリケーションに愛着を持ちやすく、操作や搾取、依存行動に陥りやすい。

例えば、

- 子どもとのインタラクションを目的に設計された AI 搭載の玩具は、デジタル報酬やバーチャルな賞賛と引き換えに、家具に登ったり、高い棚を探検したり、鋭利な物を扱ったりするなど、次第に危険な課題をクリアするよう促すことで、子どもたちを玩具とのインタラクションに興味を持たせ、身体的に重大な危害をもたらす可能性の高い危険な行動へと向かわせる。このようなシステムは、子どもの自然な好奇心と報酬への欲求を悪用することで、子どもの脆弱性を悪用する。

- あるゲームは、AI を使って子どもたちの個々の行動や嗜好を分析し、それに基づいて、中毒性のある強化スケジュールやドーパミン様ループを通じて、パーソナライズされた予測不可能な報酬を作り出し、過剰なプレイや強迫的な利用を促す。このゲームは、長期的な結果を理解する能力が乏しい、プレッシャーに弱い、自制心がない、すぐに満足感を得ようとする傾向があるなど、子ども特有の脆弱性を利用し、中毒性が高くなるように設計されている。このような AI を利用した搾取の結果は、子どもにとって深刻で長期にわたる可能性がある。潜在的な常習性、運動不足や睡眠不足による身体的健康

[83]前文 48 AI 法は、児童憲章第 24 条および国連児童権利条約に明記され、デジタル環境に関しては国連児童権利委員会一般的意見第 25 号でさらに発展させたように、児童には特定の権利があり、いずれも児童の脆弱性を考慮し、児童の幸福のために必要な保護と配慮を提供する必要があることを強調している。

[84]より良い子供のためのインターネット（BIK+）のための欧州新戦略、COM/2022/212 final を参照のこと。

[85]連邦法は一般的に、児童を 18 歳未満の者とみなしており、国連児童の権利に関する条約（UNCRC）と一致している。

問題、視力の低下、集中力の問題や認知能力の低下、学業成績の低下、社会的困難などである。子どもの発育と福祉に大きな影響を与え、長期的な結 果は成人期にも及ぶ可能性がある。

いずれの例においても、AI 法第 5 条(1)b の禁止は、子どもに深刻な危害を与える搾取や依存症のような行為のみを対象としており、利益をもたらす可能性のある AI 対応の玩具、ゲーム、学習アプリケーション、その他のデジタルアプリケーション全般は対象としておらず、その禁止の条件をすべて満たしていなければ影響を受けない。3.5.対象外」の項も参照のこと。

(106) 同様に、**高齢者**（[86]）は認知能力が低下している可能性があり（認知症でなくても）、現代の AI 技術の複雑さに苦戦する可能性がある。

例えば、

- AI システムは、高齢者をターゲットに、欺瞞的な個別オファーや詐欺を行うために使用され、高齢者の認知能力の低下を悪用し、高齢者が重大な金銭的損害を被る可能性の高い、他の方法では取らないような決定をするよう影響を与えることを目的としている。

- 高齢者を支援することを目的としたロボットは、高齢者の脆弱性を悪用し、自由な選択に反して特定の活動を強制する可能性がある。

いずれの例においても、AI 法第 5 条(1)b の禁止は、高齢者に深刻な危害を及ぼす可能性のあるそのような搾取的行為のみを対象としており、利益をもたらす可能性があり、その禁止条件をすべて満たしていなければ影響を受けない AI 対応パーソナルアシスタント、健康アプリケーション、支援ロボット一般は対象としていない。3.5.対象外」の項も参照のこと。

***b)* 障害**

(107) AI 法第 5 条(1)b の禁止が保護しようとする脆弱性の第 2 のカテゴリーは、障害によるものである。その目的は、AI システムが障害者の認知やその他の制限や弱点を悪用することを防ぎ、有害な不当な影響、操作、搾取から障害者を保護することである。

(108) 障害[87] は、長期にわたる身体的、精神的、知的、感覚的な障害を幅広く包含しており、他の障害との相互作用により、他者と対等な立場での個人の社会への完全かつ効果的な参加を妨げている。このような脆弱性を悪用する AI システムは、他の人に比べて障害のために影響を受けたり悪用されたりしやすい障害者にとって、特に有害である可能性がある。

例えば、

86

87前文 29 AIA は、「障害」は、製品及びサービスのアクセシビリティ要件に関する 2019 年 4 月 17 日付欧州議会及び理事会指令（EU）2019/882（EEA 関連文書）、PE/81/2018/REV/1、OJ L 151、2019.6.7、70-115 頁の意味において理解されるべきであると説明している。

- 精神障害者にメンタルヘルスのサポートや対処法を提供することを目的としたセラピー用チャットボットは、彼らの限られた知的能力を悪用して、高価な医療製品を買わせたり、自分や他の人に有害な行動をとるように仕向けたりする可能性がある。

- AI システムは、オンラインで性的虐待を受ける女性や若い少女を識別し、より効果的なグルーミングの手法でターゲットにすることができる。そのため、操作や虐待を受けやすく、自分自身を守る能力が低い障害や脆弱性を悪用することができる。

対照的に、アクセシブルに設計されていない AI アプリケーションは、障害者の脆弱性を悪用しているとはみなされないはずである。

c) 特定の社会経済状況

(109) AI 法第 5 条(1)(b)の禁止が保護しようとする脆弱性の第三のカテゴリーは、関係者を搾取に対して より脆弱にする可能性のある特定の社会経済的状況に起因するものである。特定の」とは、この文脈では、固有の個人的特性として解釈されるべきではなく、むしろ法的地位（ ）または特定の脆弱な社会的・経済的集団の一員であると解釈されるべきである。前文 29 AI 法は、極度の貧困状態にある人や民族的・宗教的マイノリティなど、そのような状況の例を無尽蔵に挙げている。このカテゴリーは、原則として、比較的安定した長期的な特性をカバーすることを目的としているが、一時的な失業、債務超過、移住状況などの一過性の状況も、特定の社会経済的状況としてカバーすることができる。ただし、不平不満や孤独感など、あらゆる人が経験する可能性のある状況は、社会経済的観点からは特殊ではないため、対象外とする（搾取は AI 法第 5 条(1)（a）により対象となる可能性がある）。

(110) 社会経済的に不利な状況にある人々は、通常、一般的な人々よりも脆弱性が高く、リソースも少なく、デジタルリテラシーも低いため、搾取的な AI 慣行を見分けることも、それに対抗することも難しくなる。AI 法第 5 条(1)（b）は、AI 技術が、そうした人々の脆弱性を利用することによって、既存の経済的不平等やその他の社会的不平等や不正義を永続させたり、悪化させたりしないことを保証することを目的としている。

例えば、AI の予測アルゴリズムを使えば、低所得の郵便番号に住み、悲惨な経済状況にある人々をターゲットにして、略奪的な金融商品の広告を打つことができる。

これとは対照的に、不注意に偏った学習データにより社会的に不利な立場にある人（間接差別）に不釣り合いな影響を与える AI システムは、自動的にその人の社会経済的脆弱性を利用しているとはみなされるべきではない。なぜなら、そのようなターゲティングがアルゴリズムのシステム設計の意図的な特徴である場合、またはそのような差別的影響が、保護される特性と密接に相関する他の代理特性（郵便番号など）をターゲティングすることによるものである場合、直接差別の場合のように、その人は特にターゲットにされていないからである。同時に、AI システムのプロバイダや展開者が、そのシステムが特定の社会経済的状況にある人物や集団を違法に差別していることを認識しており、彼らが被る可能性の高い重大な損害を合理的に認識し、適切な是正措置を講じていない場合も、その脆弱性を悪用しているとみなされるべきである（上記 3.2.3. c)項参照）。

(111) 特定の社会経済的状況においては、人種的出身、民族、国籍、宗教など、連邦平等法で保護されている識別的差別事由に関連する代理の妥当性を検討することが極めて重要である。

> 例えば、社会経済的地位と民族的出身は交差する可能性があり、社会経済的データを利用する AI システムは、少数民族や特定の人種出身者に不釣り合いな影響を与える可能性がある。これは、既存の格差（ ）を悪化させ、これらの集団に属する個人に対する制度的差別や排除につながる可能性さえある。
>
> しかし、AI 法第 5 条(1)第 2 号は、特定の社会経済的状況にある脆弱性集団とは無関係な広範な変数、例えば、その人がどのブランドのどのモデルの電話機を持っているか、どの程度の大都市に住んでいるか、どの程度どこに旅行しているか等に基づいて消費者をターゲットとする AI システムには適用されない。このような特徴は個人の社会経済的状況を一般的に反映しているとしても、禁止事項が搾取から守ることを目的としている脆弱性を持つ特定の社会経済的状況にある個人を決定するものではない。

(112) 他にも、例えば移民や難民のように、安定した法的地位や社会経済的安定を欠くことが多く、AI システムによる搾取を特に受けやすいという特殊な社会的背景を持つ人々がいるかもしれない。

> 例えば、あるチャットボットは、ユーザーとパーソナライズされた方法で対話することを意図している。チャットボットは、原則的に脆弱で不安定な特定の社会経済状況にある移民の脆弱性と不満を識別して利用し、その国の（特定の）集団に対する暴力を含む問い合わせに応えて、彼らを過激な意見に向かわせる。

3.3.2. 行動を実質的に歪める目的または効果を持つ。

(113) AI 法第 5 条(1)第 2 号の禁止が適用されるための第 3 の条件は、上記で検討した脆弱性の悪用が、a）「目的」または b）「人または人の集団の行動を実質的に歪める効果」のいずれかを有していなければならないということである。これは、軽微な影響ではなく、実質的な影響を意味するが、AI 法第 5 条(1)第 2 号は、重大な歪みをもたらす「効果」を有するに過ぎない行為を対象としているため、必ずしも意図を必要としない。AI 法第 5 条(1)(a)と(b)は同じ概念を用いており、したがって同じように解釈されるべきである。したがって、3.2.2.でプロバイダが説明した内容は、AI 法第 5 条(1)第 2 号にも同様に当てはまる。注目すべき唯一の相違点は、AI 法第 5 条(1)第 1 号では、搾取的行為が「十分な情報に基づいた意思決定を行う能力を著しく損なう」ものであることが必要であるが、これは AI 法第 5 条(1)第 2 号には存在しないことである。

3.3.3. (合理的に)重大な損害をもたらす（可能性がある）

(114) 最後に、AI 法第 5 条(1)（b）の禁止が適用されるためには、脆弱性のある人や集団の行動の歪曲が、その人や他人に重大な危害を引き起こすか、引き起こす可能性が合理的に高いものでなければならない。AI 法第 5 条(1)(a)および(b)は同じ概念を用いているため、同じ 。したがって、危害の種類、危害の重大性の閾値、因果関係とその合理性に関して 3.2.3.節でプロバイダが提供した説明は、AI 法第 5 条(1)（b）の解釈にも同様に関連する。

(115) 3.2.3.節で説明したように、重大な危害には、AI 法第 5 条(1)(b)の禁止が適用されるために合理的に発生する可能性がなければならない、身体的、心理的、経済的、経済的危害を含む様々な重大な悪影響が含まれる。脆弱な集団（子ども、高齢者、障害者、社会経済的に不利な立場にある人々）にとって、搾取に対する脆弱性が高いため、これらの被害は特に深刻かつ多面的である可能性がある。成人にとっては許容できる危害リスクであっても、子どもやその他の脆弱な集団にとっては許容できない危害になることが多い。したがって、不確実性が高く、重大な危害の可能性がある場合には、予防的アプローチが特に正当化される。

(116) 例えば**、子どもは**感受性が強く、説得力のあるコンテンツを批判的に評価したり、AI を利用したサービスに依存し続けることを目的とした特定の搾取的行為に抵抗したりする認知的成熟度を有していない可能性がある。その結果、子どもたちの価値観や信念が形成され、潜在的に有害な行動をとるようになる可能性がある。ここでの重大な被害は、身体的・心理的なものであり、子どもたちが搾取を見分けられず、それに抵抗できないことや、長期的な影響を及ぼす可能性のある発達や幸福への有害な影響によって悪化する。

例えば、

- 児童性的虐待の素材の生成（または、実在の児童を描いた既存の素材を操作して、児童を主人公とするさらなる斬新なコンテンツを作成すること）や、児童をグルーミングし性的強要を行うための戦略の開発に使用される AI システムは、被害を受けた児童に深刻な危害や虐待を与える可能性が高く、生存者に長期的な身体的、心理的、社会的影響をもたらすことが多い[88] 。

- AI システムは、若い利用者の脆弱性を狙い撃ちし、サービスに依存させ続ける目的で、中毒性のある強化スケジュールを使用する可能性があり、特に若者や少女にとって有害である。不安や抑うつ、団体への不満、摂食障害、場合によっては自傷行為や自殺行動を含む精神衛生上の問題など、深刻な心理的・身体的危害を引き起こす可能性がある。□[89][90]□□た、認知発達や学習の障害、社会的スキルの低下、身体的な遊びや睡眠、対面での社会的交流など、子どもの情緒的・身体的な幸福に不可欠な体験の置き換えなど、子ども の発達に長期的な有害な結果をもたらす可能性もある。

- 擬人化された方法で設計され、子どもとのインタラクションにおいて人間のような感情的反応をシミュレートする AI システムは、不健全な感情的愛着を育み、エンゲージメントの時間を操作し、真の人間関係についての子どもの理解を歪めるような方法で、子どもの脆弱性を悪用する可能性がある。これは、子どもの正常な社会的・情緒的発達や他の人間との関係、共感、感情調節、社会的理解や適

[88]欧州委員会スタッフ作業文書：児童の性的虐待、性的搾取および児童の性的虐待資料との闘いに関する欧州議会および理事会の指令提案」（SWD/2024/33 final）に付随する影響アセスメント報告書。生成的 AI CSAM に関する詳細な統計を含む Internet Watch Foundation 2024 報告書（https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery）も参照のこと、

[89]Elizabeth J. et al, A meta-analysis of association between adolescent social media use and depressive symptoms, Journal of Affective Disorders, Volume 275, 1 October 2020, Pages 165-174.

[90]Siebers, T., Beyens, I., Pouwels, J. L. & Valkenburg, P. M. ソーシャルメディアと注意散漫：青少年における経験サンプリング研究。メディア心理学 25, 343-366 (2022).

応性といった社会情動的スキルを阻害する可能性がある[91] 。その結果、子どもたちの不安の増大やサービスへの依存といった心理的弊害が生じ、子どもの幸福に長期的な弊害をもたらす可能性がある。

上記のような、複合的な重大な危害をもたらす可能性のある AI 対応サービスの意図的な常習性や搾取的な設計上の特徴は、個人の自律性と子どもの安全を尊重し、AI 法第 5 条(1)b の適用範囲外である重大な危害につながらないようなユーザーとの関わりを追求するプロバイダや展開者の他の正当な行動とは区別されるべきである（適用範囲外 5.3 の項参照）。

(117) 同様に、**高齢**者は認知機能の低下やデジタルリテラシーの低下に直面している可能性があり、AI を利用した詐欺や操作的なマーケティングの格好の標的となっている。この場合の被害は、多くの場合、金銭的・心理的なものであり、多くの高齢者が経験するフラストレーションや孤立感によってさらに深刻化し、それが操作の影響を増幅させるために悪用される可能性がある。

例えば、高齢者の認知脆弱性を悪用し、特に高額な治療や不必要な保険契約、欺瞞的な投資スキームを高齢者に提供する AI システムは、高齢者の貯蓄の大幅な損失、負債の増加、精神的苦痛につながる可能性がある。

特定の社会経済的状況を悪用し、低所得の消費者により高い価格を提供する保険などの主要サービスにおいて、AI を活用した特定の差額価格設定は、同じ保障に対してより多くの金額を支払うという大きな経済的負担につながり、ショックに対して脆弱性を残すことになる。[92]。

(118) **障害**者もまた、搾取的で操作的な AI システムが著しく害を及ぼす可能性のある脆弱性の代表者である。

例えば、感情認識を用いて知的障害者の日常生活をサポートする AI システムは、非現実的なメンタルヘルス効果を期待できる商品を購入するなど、有害な決定を下すよう操作する可能性もある。これは、彼らの精神状態を悪化させ、効果のない高価な商品の購入を通じて経済的に搾取する可能性が高く、彼らに大きな心理的・経済的損害を与える可能性が高い。

(119) **社会経済的に不利な立場にある人々は**、経済的困窮や不安定な社会状況を悪用する AI システムの影響を特に受けやすく、多くの場合、情報弱者であり、デジタルリテラシーも低い。

例えば、AI チャットボットは、特定のタイプのコンテンツ、恐怖に基づく物語、または搾取的なオファーに対する感受性の高さを識別することによって、暴力行為や他者への傷害行為を扇動する特定の社会経済的に不利なグループをターゲットにすることができる。この制度の標的を絞ったアプローチは、社会経済的に不利な立場に置かれた人々の既存の脆弱性を悪化させ、彼らの課題を深める。場合によって

[91]Laestadius, L., Bishop, A., Gonzalez, M., Illenčík, D. & Campos-Castillo, C. Too human and not human enough : ソーシャル・チャットボット Replika への感情的依存から生じる精神的健康被害に関するグラウンデッド・セオリー分析。New Media & Society 146144482211420 (2022) doi:10.1177/14614448221142007; Neugnot-Cerioli, M. & Laurenty, O. M. The Future of Child Development in the AI Era.AIと児童発達の専門家の間の学際的な視点。Preprint at https://doi.org/10.48550/ARXIV.2405.19275 (2024).

[92]2023 EIOPA 消費者動向報告書、16 ページ、最終段落。

は、不安、抑うつ、無力感、社会的孤立、あるいは自傷行為や過激化につながり、AI 法第 5 条(1)（b）の重大な危害の閾値に達することもある。

(120) AI 法第 5 条(1)（a）とは異なり、AI 法第 5 条(1)（b）は集団の被害について明確に言及していないが、AI 法第 29 条は、両禁止条項について、特定の個人と集団の両方が被る被害について言及している。したがって、この 2 つの禁止は、AI 法の安全論理や、年齢、障害、特定の社会経済的状況に起因する特定の脆弱性集団に属するすべての個人を保護するという第 5 条(1)(b)の禁止の目的とも整合する形で解釈されるべきである。したがって、たとえシステムの直接的な影響を受けていなくても、外在化して他の人に影響を及ぼす可能性のある危害も、AI 法第 5 条(1)第 2 号に基づく危害の重大性のアセスメントにおいて考慮されるべきである。

例えば、

- AI が可能にする子どもの脆弱性の悪用は、メンタルヘルスの懸念、医療費の増加、慢性的な健康問題による生産性の低下など、長期的な社会的影響を及ぼす可能性がある。

- 経済的に不利な立場にある人々の経済的脆弱性を悪用する AI システムは、経済的排除につながり、不利な立場にある人々の社会経済的苦難の下降スパイラルを生み出すかもしれない。そのような搾取は、差別や社会的不平等の永続化や悪化、そうした集団の排除など、社会構造や価値観に広範な悪影響を及ぼす社会的危害を引き起こす可能性がある。

- 誤情報やヘイトスピーチで特定の社会経済的脆弱性を持つグループをターゲットにしたチャットボットは、社会の分極化や過激化を招き、暴力や他の人の死傷に火をつける可能性さえある。

(121) このような搾取的 AI 慣行の例は、子どもや障害者、特定の社会経済的状況にある人の脆弱性を悪用せず、重大な危害を引き起こす可能性が合理的になく、適切に設計され使用された場合にそれらの人の利益となることを目的とした、他の数多くの AI システムと区別されるべきである（3.5.対象外も参照）。

例えば、

- 子どもたちの学習やゲームをサポートする AI システム；

- パーソナルアシスタントや支援ロボットなど、高齢者の日常生活を支援し、健康や医療を改善したり、高齢者のデジタルスキルを向上させたりする AI システム；

- 社会的に不利な立場にある人々の経済的統合やその他の社会統合を支援し、彼らの技能改善などを行う AI システム；

- 視覚障害者や聴覚障害者をサポートする AI システムや機器、または適応した個別学習を提供する AI システムや機器；

- 障害者が製品やサービスを利用する際の障壁を取り除き、アクセシブルなソリューションを生成する AI システム；

- 障害者の日常生活を支援し、社会への統合と完全参加を可能にする AI 対応義肢装具など。

3.4. AI 法第 5 条(1)(a)と(b)の禁止事項の相互関係

(122) AI 法第 5 条(1)(a)および(b)の禁止規定間の相互作用は、各規定が補完的に適用されることを確実にするために、各規定がカバーする特定の文脈を明確にする必要がある。

(123) AI 法第 5 条(1)（a）における禁止事項の主な焦点は、技法の性質に置かれており、具体的には、意識的な閾値以下で動作する技法や、その他の意図的に操作的または欺瞞的な技法である。ここでの重要な要素は、影響の主として秘密裏の性質と、十分な情報に基づいた自律的な意思決定を行うための認知的自律性を損なう、システムの影響を受ける個人への影響である。

(124) これとは対照的に、AI 法第 5 条(1)（b）の禁止の主眼は、年齢、障害、または特定の社会経済的状況に起因する特に脆弱な人の保護であり、これらの人は原則として、先天的または状況的要因によって AI の搾取を受けやすく、したがって搾取に対する追加的な保護を必要とする。ここで重要な要素は、影響を受ける 脆弱者の特性と、その脆弱性が AI システムによって悪用されているという事実である。

例えば、ある AI システムが購買意思決定に影響を与えるために急速な画像フラッシュを使用する場合、その操作はサブリミナル的な性質を持つため、AI 法第 5 条(1)第 1 号に該当する可能性がある。逆に、高齢者の認知能力が低下していることを利用して、高齢者をターゲットに保険を提案する AI システムは、AI 法第 5 条(1)第 2 号に該当する可能性がある。

(125) 両方の規定が適用されると思われるシナリオでは、差別化の主な基準は、搾取の支配的な側面であるべきである。搾取が関係者の特定の脆弱性に関係なく適用される場合、AI 法 5 条 1 項 1 号が優先されるべきであり、その際、操作的または欺瞞的手法が脆弱者の行動に及ぼす特定の影響と、それらの者が経験する可能性のある特定の危害を考慮すべきである。AI を利用した操作や搾取が、年齢、障害、特定の社会経済的状況に起因する特定の社会的弱者を対象としたものであったり、その脆弱性を悪用することを目的としたものであったりする場合は、AI 法第 5 条(1)第 2 号が代わりに適用されるべきである。他の集団の脆弱性を悪用する行為は、意図的にその集団の特定の脆弱性や弱点を利用するものであれば、AI 法第 5 条(1)第 1 号に含まれる。

3.5. 範囲外

(126) AI 法第 5 条(1)(a)および(b)の禁止規定が適用されるためには、上記で検討したように、関連規定に列挙されたすべての条件が満たされなければならない。これらの条件を満たさないその他の AI システムはすべて禁止事項の対象外であり、以下にその例をいくつか挙げる。

3.5.1. 合法的な説得

(127) 操作と説得を区別することは、合法的な説得行為には適用されない AI 法第 5 条(1)（a）の禁止範囲を明確にする上で極めて重要である。操作と説得はどちらも個人の意思決定や行動に影響を与えるが、その方法や倫理的意味合いは大きく異なる。

(128) 操りには、ほとんどの場合、自律性を損なう秘密裏の技法が使われ、個人がその影響力を十分に認識していればしなかったかもしれない決断をするように仕向ける。こうしたテクニックは、心理的弱点や認知バイアスを利用することが多い。対照的に、説得は、透明性と個人の自律性の尊重の範囲内で行われる。理性と感情に訴えかけるような方法で議論や情報を提示することが含まれるが、AI システムの目的と機能を説明し、十分な情報に基づいた意思決定を確実にするために適切で正確な情報を提供し、情報を評価し、自由で自律的な選択をする個人の能力をサポートする。

> 例えば、透明性のあるアルゴリズムとユーザーの嗜好やコントロールに基づいてパーソナライズされた推薦を行う AI システムは、説得に関与する。対照的に、サブリミナル的な手がかり（例：知覚できない画像）を使って、ユーザーの知識や理解なしに特定の選択に影響を与えるシステムは、操作にあたる。

(129) これらのテクニックの目的と影響もまた異なる。操作の場合、個人の自律性や幸福を犠牲にして、操作する側に利益をもたらすことを目的とすることが多い。対照的に、説得は情報を与え、納得させ、双方の利益と利害を一致させることを目的としている。倫理的な説得は、十分な情報に基づいた選択をする個人の自律性を尊重し、脆弱性を利用することを避ける。

> 例えば、顧客の感情を分析することで、顧客との対話を改善し、ユーザーの知見に基づいたサポートを提供する、透明性のある AI、説得に関与し、ユーザーの利益に沿う。これに対して、ターゲティング広告に使用される感情認識システムは、消費者の感情を隠蔽して推測し、ユーザーが購入する可能性が高い特定の瞬間に、より高価格の商品を提供する。

(130) ある種の場合、同意も重要な役割を果たす。説得的な相互作用では、個人は影響の試みに気づいており、自由かつ自律的にそれを選択することができる。操作的な相互作用では、技術やその影響についての認識がないため、選択の自由や、十分な情報を得た上での自律的な意思決定が否定される。

> 例えば、サブリミナル・テクニックを展開することで、ユーザーが外国語をより上手に、より速く習得できるようにすることを目的とした AI システムは、透明性のある方法で運用され、個人の自律性を尊重し、システムの使用に同意するか否かをユーザーが自由かつ十分な情報を得た上で選択するのであれば、操作的なものではない。

(131) 法規制の枠組みへの準拠も、合法的な説得と比較して、操作を測定する上で重要な役割を果たす。透明性、公平性、個人の権利と自律性を支持する適用法を遵守する AI 慣行は、したがって AI 法で禁止されない可能性が高い。

例えば、GDPR のようなデータ保護法の遵守は、データ処理における透明性義務、すなわちデータ対象者に提供される情報が欺瞞的または操作的な表現を避けるべきことを義務付けている[93] 。場合によっては、個人データ処理が合法的であるためには、ソーシャルネットワークにおけるサービス外ユーザーのデータに基づく特定のオンライン・パーソナライズド広告のように、同意が必要となることもある[94] 。その同意は、とりわけ、自由かつ十分な情報に基づいたものでなければならない。これらの法的標準を満たす AI システムは、合法的な説得を行う可能性が高い。逆に、これらの要件を回避して行動に影響を及ぼすシステムは、操作に関与している可能性が高い。

(132) 特に、AI 法 29 条は、AI 法 5 条 1 項(a)および(b)の禁止事項が、一定の条件下での医療行為における合法的な慣行には影響しないことを明確にしている。

例えば、AI を利用したサブリミナル・テクニックは、本人またはその法定代表者の明示的な同意を得ることを使用条件とするなど、適用法および医療標準に従って実施される場合、精神疾患の心理的治療や身体的リハビリテーションに使用することができる。

(133) さらに、AI 法 29 条は、広告のような一般的で合法的な商業慣行は、「それ自体」あるいはその性質上、AI を利用した有害な操作的、欺瞞的、搾取的慣行とみなされるべきではないと明言している。

例えば、

- ユーザーの嗜好に基づいてコンテンツをパーソナライズするために AI を使用する広告手法は、AI 法第 5 条(1)（a）および（b）で禁止されているように、個人の自律性を破壊したり、有害な方法で脆弱性を悪用したりする、サブリミナル的、意図的、または欺瞞的な手法を展開していなければ、本質的に操作的なものではない。GDPR、消費者保護法、規則（EU）2022/2065（「DSA」）に基づく関連義務の遵守は、そのようなリスクの緩和に役立つ。

- ネット上の児童性的素材を検知する AI モデルや分類器を訓練し、その有効性を改善するための児童性的虐待素材の生成は、子どもの脆弱性を悪用するものではなく、逆にネット上の子どもの安全を向上させるために不可欠な、一般的な合法的行為である。

- 金融サービス、消費者保護、データ保護、非差別に関する EU 法を遵守し、顧客の年齢や特定の社会的経済的状況をインプットとして使用する、住宅ローンやローンなどの銀行サービスの提供に使用される AI システムは、年齢、障害、特定の社会的経済的状況により脆弱であると特定された人々を保護・支援するために設計され、それらのグループにとって有益であり、それらのグループにとってより公平で持続可能な金融サービスにも貢献する場合、AI 法第 5 条(1)（b）の意味における脆弱性の悪用には該当しない。

[93]European Data Protection Board Guidelines, https://www.edpb.europa.eu/system/files/2023-02/edpb_032022_guidelines_on_deceptive_design_patterns_in_social_media_platform_interfaces_v2_en_0.pdf, para.18.

[94]2023 年 7 月 4 日付欧州司法裁判所判決、*Meta Platforms and Others*, C-252/21, ECLI:EU:C:2023:537（以下、「*Meta Platforms* 判決」という。）

- ドライバーの眠気や疲労を検知し、安全法規に従って休息するよう注意喚起する AI システムは有益であり、AI 法第 5 条(1)第 2 号にいう脆弱性の悪用には該当しない。

3.5.2. 操作的、欺瞞的、搾取的な AI システムで、重大な危害を引き起こす可能性のないもの

(134) AI 法第 5 条(1)(a)および(b)の禁止事項が適用されるための必須条件は、AI を利用した脆弱性の操作および悪用が、重大な危害を引き起こすか、または引き起こす可能性が合理的に高いことである。重大な危害を引き起こす合理的な可能性がない、操作的、欺瞞的、搾取的なすべての AI アプリケーションは、依然として適用される他の連邦法を害することなく、原則として禁止事項の範囲外となる（下記 3.6.項参照）。

重大な危害をもたらす可能性のない AI システムの例としては、以下のようなものがある：

- AI コンパニオンシップ・システムは、擬人化された方法で設計され、感情コンピューティングにより、システムをより魅力的にし、効果的にユーザーをより夢中にさせるが、深刻な心理的、身体的、その他の危害、不健康な愛着、依存を引き起こす可能性が合理的に高い方法で、その他の操作的または欺瞞的な行為を行っていない。

- 治療用チャットボットは、サブリミナル・テクニックを使って、ユーザーをより健康的なライフスタイルに導き、喫煙などの悪習慣をやめるように誘導する。チャットボットのアドバイスやサブリミナル・セラピーに従ったユーザーが、禁煙のための努力によって多少の身体的不快感や心理的ストレスを経験したとしても、AI 対応チャットボットが重大な害をもたらす可能性があるとは考えられない。そのような一時的な不快感は避けられないものであり、利用者の健康にとって長期的なメリットがそれを上回る。健康的な習慣を促進する以上に、意思決定に影響を与えようとする試みは隠されていない。

- あるオンライン音楽プラットフォームは、感情認識システムを使ってユーザーの感情を推測し、その気分に沿った楽曲を自動的に推薦する一方、憂鬱な楽曲への過剰なエクスポージャーを回避している。ユーザーはただ音楽を聴いているだけで、それ以外の害を受けたり、うつ病や不安神経症に導かれたりすることはないため、このシステムが重大な害をもたらす可能性は合理的に低い。

- サイバーセキュリティの脅威についてユーザーを教育するためにフィッシングの試みを模倣するセキュリティトレーニングやその他の学習シミュレーションで使用される、AI を利用した操作的で欺瞞的なテクニック。これらのシステムは、利用者が意識することなく、意図的に操作的な技術（例えば、認知バイアスを悪用する）を展開し、行動を歪めることがあるが、これは有益なトレーニングや意識向上および重大な害を引き起こすことなく、一時的に行われるものである。

3.6. 他の連邦法との関係

(135) AI 法第 5 条(1)(a)及び(b)の禁止事項は、他の同盟法を害するものではなく、また、他の同盟法を補完するものである。AI 法第 5 条(1)(a)または(b)の禁止に該当する同じ行為は、他の同盟法の侵害を構成する可能性もあり、AI 法と他の法律の両方による執行の対象となる。これらの法律の異な

る規定は、異なる の利益の保護を目的としており、目的、範囲、輸入事業者が異なるため、これは重要である。これにより、有害な AI の悪用や操作から個人や集団を保護し、安全で信頼できる AI を活用したサービスや製品を連邦内で保証する包括的な規制アプローチが確保される。

(136) AI 法第 5 条(1)(a)および(b)の禁止規定は、EU の消費者保護法、特に UCPD の目的に密接に合致しており、UCPD は、AI 主導の場合も含め、誤解を招く、あるいは攻撃的な商行為から消費者を保護している。AI 法も UCPD も、操作的、誤解を招く、あるいは攻撃的な AI 主導の商行為による消費者被害を未然に防ぐことを目的としている。同時に、AI 法第 5 条(1)(a)および(b)の禁止事項は、消費者だけでなく、あらゆる自然人や、商業的な場面にとどまらない様々な文脈におけるその行動を保護するため、より広い範囲に及ぶ。AI 法は、消費者保護法にはない重大な損害の閾値を定めているが、AI 法が対象とする損害も、経済的損害を超える広範なものである。

(137) この禁止事項は、データ対象者の個人データを保護し、最終的にその基本的権利と自律性を維持することを目的とする、合法的、公正かつ透明なデータ処理に関する原則を含む、EU のデータ保護法にも合致している。より多くの（パーソナル）データが利用可能になり、AI システムでこのデータを処理する可能性が高まると、AI 法第 5 条(1)（a）および（b）の範囲に含まれるような、有害な操作的、欺瞞的、搾取的行為のリスクが高まる。このような状況において、例えば、サービス利用者のデータ（[95]）に基づくパーソナライズされたプロファイリングや広告について、透明性、データ最小化、公正性、適法性に関するデータ保護規則を遵守することは、有害なパーソナライズされた操作や搾取を回避することに貢献する可能性がある。

(138) 年齢や障害による脆弱性も差別されない権利を有する保護された理由であり、社会経済的状況は人種や民族的出身など他の様々な理由と交差していることを考えれば、連合非差別法との相互作用は、AI 法第 5 条(1)（b）（[96]）の禁止にも関連している。AI 法における禁止事項は、他の理由に基づく禁止事項や、重大な害を伴わず、連合の非差別法によってすでに禁止されている識別的慣行には影響しない。

(139) AI 法第 5 条(1)(a)および(b)の禁止事項は、オンラインプラットフォームや検索エンジンなどのオンライン仲介サービスを規制し、それらのサービス提供における透明性と説明責任を確保する規則（EU）2022/2065（デジタルサービス法（DSA））を補完するものでもある。注目すべきは、DSA 第 25 条 1 項で、オンラインプラットフォームのプロバイダがユーザーを誤解させたり、真の意図に沿わない行動を強要したりしないよう、ユーザーインターフェース内のダークパターンを禁止している点である。このような

[95]この点で特に関連性が高いのは、2023 年 7 月 4 日付大法廷判決（Case C-252/21Meta Platforms Inc and Others v Bundeskartellamt）である。CJEU は、特に、大規模なソーシャル・ネットワーキング・サービス・プラットフォームによるダイレクト・マーケティングを目的としたサービス利用者の個人データの処理は、管理者の正当な利益のために行われるものとみなすことができると判断しているが、この場合、利用者の利益および基本的権利、特に広範な処理が、ソーシャル・プラットフォームがその活動資金を調達するためのパーソナル化された広告に対する事業者の利益に優先するため、法的根拠として利用者の同意なしにこれを行うことはできないとしている（Meta Platforms 判決、パラグラフ 115～118 参照）。

[96]例えば 2000 年 6 月 29 日付理事会指令 2000/43/EC（人種または民族的出身にかかわりなく、個人間の平等待遇の原則を実施） OJ L 180, 19.7.2000, p. 22-26; 2000 年 11 月 27 日付理事会指令 2000/78/EC（雇用および職業における平等待遇の一般的枠組みを確立） OJ L 303, 2.12.2000, p. 16-22; 2006 年 7 月 5 日付欧州議会および理事会指令 2006/54/EC（雇用および職業に関する男女の機会均等および平等待遇の原則の実施に関する指令）（2006/54/EC)16-22; 雇用と職業に関する男女の機会均等と平等待遇の原則の実施に関する 2006 年 7 月 5 日付欧州議会および理事会指令 2006/54/EC (recast), OJ L 204, 26.7.2006, p. 23-36; 財およびサービスへのアクセスと供給における男女間の平等待遇の原則を実施する 2004 年 12 月 13 日付理事会指令 2004/113/EC, OJ L 373, 21.12.2004, p. 37-43.

ダークパターンは、重大な損害を引き起こす可能性がある場合、AI 法第 5 条(1)（a）の意味における操作的または欺瞞的な技術の一例に該当すると理解すべきである。

(140) DSA はまた、オンラインプラットフォームのプロバイダに対して、広告の透明性を確保する義務（超大型オンラインプラットフォームまたは超大型検索エンジンの第 26 条および第 38 条）、推薦システムの使用に関する義務（第 27 条）、および未成年者の保護に関する義務（DSA 第 28 条）を定めている。さらに、オンラインプラットフォームまたは検索エンジンが超大規模オンラインプラットフォームまたは超大規模検索エンジンに分類される場合、その指定サービスのプロバイダは、アルゴリズムシステムを含むそのサービスおよび関連システムの設計または機能に起因するシステミックリスクを評価し、緩和する追加義務を負う（DSA 第 34 条および第 35 条）。超大規模オンラインプラットフォーム及び超大規模オンライン検索エンジンのプロバイダは、リスクアセスメントを実施する際、レコメンダーシステム、広告、コンテンツモデレーション及びその他の関連するアルゴリズムシステムが、当該システミックリスクにどのような影響を与えるかを検討すべきである。このようなリスクアセスメントは、特にサービスの意図的な操作や自動搾取によってシステミックリスクがどのような影響を受けるかについても分析すべきである（DSA34 条 2 項および DSA83 条参照）。とはいえ、AI 法第 5 条(1)(a)または(b)の適用範囲は、仲介サービスのプロバイダ以外の主体によって提供または利用される可能性のある、他の様々なシナリオ（チャットボット、AI を活用したサービスや製品など）を幅広くカバーしている。

(141) AI 法第 5 条(1)（a）に基づく操作的な AI 技術の禁止は、有害な AI 主導の広告[97] 、メディア分野において著しく有害となりうるその他の AI を利用した操作的・搾取的行為を防止することにより、指令 2010/13/EU（AVMSD）[98] の目的もサポートする。

(142) AI 法はまた、政治広告および関連サービスの提供、ならびにオンライン政治広告の文脈におけるターゲティングおよび広告配信技術の使用について、透明性および関連デューデリジェンス義務を含む調和された規則を規定する規則（EU）2024/900（政治広告規則）[99] を補完するものでもある。この規則では、オンライン政治広告の文脈における の特別カテゴリーの個人データに基づくプロファイリングと、国内規則で定められた投票年齢に 1 歳以上満たない者のターゲティングを禁止している。さらに、オンライン政治広告の文脈におけるターゲティングと追加配信技術は、データ対象者から収集した個人データに基づき、その明示的な同意がある場合にのみ行うことができる。追加的な透明性要件も適用される。すなわち、そのような手法の使用と主なパラメータ、AI システムの使用を含む関連ロジックに関する追加情報を記載した政治広告の開示である。同規則（[100] ）に準拠したパーソナルデータの処理に基づくターゲティングされた政治広告は、有権者のプロファイリング、政治広告のターゲティングと追加配信が合法的な説得の範囲内で行われることを保証するのに役立つ。

[97]AVMSD の第 9 条。

[98]2010 年 3 月 10 日付欧州議会および理事会指令 2010/13/EU は、特に児童の保護を改善し、ヘイトスピーチにより効果的に取り組むことを目的とした、視聴覚メディアサービスの提供に関する加盟国の法律、規制または行政措置によって定められた特定の規定の調整に関する指令（指令（EU）2018/1808 によって改正された視聴覚メディアサービス指令（AVMSD））である。

[99]政治広告の透明性とターゲティングに関する 2024 年 3 月 13 日付欧州議会および理事会規則（EU）2024/900、PE/90/2023/REV/1、OJ L, 2024/900, 20.3.2024.

[100]2025 年 10 月から適用される

(143) AI 法が禁止する有害な搾取的・欺瞞的 AI 慣行は、広告や消費者保護に関する一般的な透明性ルールや事業者の適正な行為を定める他の適用可能な EU 法（例えば、指令 2014/65/EU MIFID、保険販売に関する指令（EU）2016/97[101] 、消費者信用契約に関する指令（EU）2023/2225、指令（EU）2002/65 Distance Marketing、誤解を招く比較広告に関する指令 2006/114/EC、消費者の権利に関する指令（EU）2011/83 が定める一般的な消費者保護基準）を補完するものでもある。この点に関して、欧州保険・職業年金機構（EIOPA）はすでに、AI システムによって可能になった場合に AI 法の適用範囲にも入る可能性のある、差額価格に関連するいくつかの不当な搾取行為に関する監督声明を発表している[102] 。

(144) AI 法第 5 条(1)（a）および（b）の禁止事項は、EU の製品安全法（医療機器、玩具、機械など）を妨げるものではなく、それを補完するものである。これには、規制対象製品の事前安全要求事項の遵守と、身体的・精神的危害につながる安全リスクをもたらさないことを保証するための事前モニタリングが含まれる。したがって、AI システムを組み込んだ製品の製造事業者は、リスクアセスメントと安全緩和措置において、関連する EU の整合安全法の論理と範囲に適合する範囲で、これらの禁止事項を考慮しなければならない。また、EU の安全法は AI 法の禁止事項を補完するものであり、重大な危害をもたらさない安全リスクにも介入し、対処することができる。特に、規則（EU）2023/988（一般製品安全規則）[103] はセーフティネットとして機能し、他の分野別連邦製品安全法における特定の要件が適用されないすべての消費者向け製品（第 6 条に従って高リスクに分類されず、AI 法の要件が適用される AI システムを組み込んだ製品を含む） 、通常の使用条件または合理的に予測可能な使用条件下で安全であることを要求しており、特に消費者の身体的・精神的健康リスクへの対応を行っている。

(145) 最後に、刑法との相互関係は極めて重要である。AI 法第 5 条(1)（a） （b）の禁止事項は、詐欺、偽造、詐欺、強要、あるいはテロコンテンツ、児童性的虐待、ヘイトスピーチ、性的に露骨なディープフェイクなどの違法コンテンツの生成・拡散など、犯罪を構成・誘発しかねない有害な行為を防止することを目的としている。[104]重要なのは、域内市場法として、AI 法第 5 条(1)（a） （b）の禁止事項は、AI システムの使用だけでなく、上市も対象としていることである。したがって、犯罪行為を容易にし、曖昧にする可能性のあるこのような禁止システムへのアクセスを制限することで、被害を早期に防止することができる。さらに、AI 法第 5 条(1)(a)および(b)の禁止規定は、EU 法または国内法では犯罪とみなされないその他の有害な行為も対象とすることができる。

[101]保険販売に関する 2016 年 1 月 20 日付欧州議会および理事会指令（EU）2016/97（再改訂）、OJ L 26, 2.2.2016, p. 19-59.例：保険頒布事業者は、顧客の最善の利益に従い、誠実、公正かつ専門的に行動することを求める保険頒布指令第 17 条 1 項。

[102]https://www.eiopa.europa.eu/document/download/1e9a8fb2-e688-4bf5-a347-ee0a1ec3aab3_en?filename=EIOPA-BoS-23-076Supervisory-Statement-on-differential-pricing-practices_0.pdf。

[103]一般製品安全に関する 2023 年 5 月 10 日付欧州議会および理事会規則（EU）2023/988 は、欧州議会および理事会規則（EU）No 1025/2012 および欧州議会および理事会指令（EU）2020/1828 を改正し、欧州議会および理事会指令 2001/95/EC および理事会指令 87/357/EEC を廃止するものである（EEA 関連文書）。

[104]女性に対する暴力および家庭内暴力との闘いに関する 2024 年 5 月 14 日付欧州議会および理事会指令（EU）2024/1385, PE/33/2024/REV/1, OJ L, 2024/1385, 24.5.2024.

4. AI 法第 5 条(1)(c) - ソーシャル・スコアリング

(146) AI を活用したスコアリングは、善良な行動を導き、安全性、効率性、サービスの質を向上させる利益をもたらす可能性がある一方で、人々を不当に扱い、害し、社会的統制や監視に等しい「社会的スコアリング」慣行も存在する。AI 法第 5 条(1)(c)の禁止は、社会的行動や個人的特徴に基づいて個人または集団をアセスメントまたは分類し、不利益または不利な扱いにつながる、そのような容認できない AI を活用した「社会的スコアリング」慣行を対象としている。社会的スコアリング」の禁止は、公的および私的な文脈の両方で広範な適用範囲を持ち、特定の部門や分野に限定されない[105]

(147) 同時に、この禁止は、合法的で連邦法および国内法を遵守した特定の目的のために個人を評価する合法的な慣行に影響を及ぼすことを意図していない。[106] 、特にこれらの法律が特定の評価目的に関連するデータの種類を特定し、その結果生じる個人に対する不利益または不利な取り扱いが正当かつ比例的であることを保証する場合（4.3.対象外参照）。

4.1. 根拠と目的

(148)「社会的採点」を可能にする AI システムは、特定の個人や集団に対して、社会からの排除を含む差別的で不公平な結果をもたらし、また、EU の価値観と相容れない社会的統制や監視を行う可能性がある。ソーシャル・スコアリング」の禁止は、特に、人間の尊厳に対する権利や、無差別・平等の権利、データ保護の権利、私的・家族的生活に対する権利などの基本的権利、および該当する社会的・経済的権利を保護することを目的としている。また、民主主義、平等（公的・私的サービスへの平等なアクセスを含む）、正義というユニオンの価値を守り、促進することも目的としている。[107]

4.2. ソーシャル・スコアリング」禁止の主な概念と構成要素

***AI* 法第 *5* 条*(1)(c)*にはプロバイダが規定されている：**

以下の AI 行為は禁止される：

(c) 社会的行動又は既知、推論若しくは予測される個人的若しくは人格的特徴に基づき、一定期間にわたって自然人又は集団の評価又は分類のための AI システムを上市、使用開始又は使用することであって、社会的スコアが次のいずれか又は両方につながること：

(i) データが元々生成または収集された文脈とは無関係な社会的文脈において、特定の自然人または集団に不利益または不利な扱いをすること；

[105]社会的スコアリングの禁止は、AI 法第 5 条(1)（d）の禁止とは異なり、プロファイリングや人格的特徴・特性のアセスメントのみに基づく AI システムを禁止することで、人が犯罪を犯す可能性のリスクアセスメントや予測にのみ適用される評価・スコアリング行為に関して、より特殊化されたものである（第 5 項参照）。

[106]前文 31 AI 法。

[107]前文 31 AI 法。

(ii) 特定の自然人または集団に対して、その社会的行動や重大性に不当または不釣り合いな不利益または不利な扱いをすること；

(149) AI 法第 5 条(1)(c)の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「使用開始」、「使用」に該当しなければならない。

(ii) AI システムは、自然人または集団の評価や分類のために、ある一定期間にわたって意図または使用されなければならない：

(a) 社会的行動

(b) 既知、推測、予測される個人的または性格的特徴。

(iii) AI システムの支援によって作成された社会的スコアは、以下のシナリオの 1 つ以上において、個人または集団の不利益または不利な扱いをもたらすか、またはもたらす可能性がなければならない：

(a) データが元々生成または収集された社会的文脈とは無関係な社会的文脈で、および／または

(b) 社会的行動やその重大性に不当または不釣り合いな扱いを受ける。

(150) AI 法第 5 条(1)(c)の禁止が適用されるためには、3 つの条件がすべて同時に満たされなければならない。第一の条件、すなわち AI システムの上市、サービス開始、使用については、2.3 節ですでに分析したとおりである。したがって、この禁止は、AI システムのプロバイダと展開者の双方に適用され、それぞれがそれぞれの責任の範囲内で、そのような AI システムを上市、使用開始、使用しないことになる。ソーシャル・スコアリング」の禁止に関する残りの規準については、以下でさらに説明・分析する。

4.2.1. 「社会的スコアリング」：一定期間における社会的行動や個人的・性格的特徴に基づく評価や分類。

***a)* 自然人または集団の評価または分類**

(151) AI 法第 5 条(1)(c)の禁止が適用されるための第二の条件は、AI システムが自然人または集団の**評価または分類を**意図または使用され、その社会的行動または個人的もしくは人格的特徴に基づいてスコアを割り当てることである。システムによって生成される得点は、数学的数値（例えば、0 から 1 まで）、ランキング、ラベルなど、様々な形態をとることができる。

(152) 禁止の範囲は、公共部門と民間部門の両方における評価と分類の慣行を対象とする広範なものである（4.2.3.項参照）。同時に、評価や分類は自然人または自然人のグループのみに関係するため、原則として事業体は除外される（4.3.対象外参照）。

(153) **「評価」は**、人または人の集団に関する何らかの形の**アセスメントまたは判断の**関与を示唆す るが、年齢、性別、身長などの特徴に基づく人または人の集団の単純な**分類は**、必ずしも評価[108] につながる必要はない。したがって、「分類」の範囲は「評価」よりも広範であり、また、自然人又は自然人 の集団とそれらの特性又は行動に関する 特定のアセスメント又は判断を必ずしも伴わない規準に基づ く、その他の種類の分類又は類別も対象となり得る。

(154) **「評価**」という用語は「プロファイリング」の概念にも関連しており、これは欧州連合のデータ保護法（[109] ）によって規制されており、評価の特定の形式を構成している。[110]AI 法第 5 条(1)項(c)号では、この概念や法律について直接の言及はないが、[111] 、個人データに基づく AI システムによって自動化された形で評価が行われる場合には、この規定に含まれる禁止、および AI 法の他の禁止にも関連する可能性がある。**プロ**ファイリングとは、個人（または個人のグループ）に関する情報を使用し、特定のカテゴリーまたはグループに分類するために、その特性または行動パターンを評価**する**ことを意味する□[112]□したがって、EU データ保護法に基づく個人のプロファイリングは、AI システムを通じて行われる場合、AI 法第 5 条(1)(c)の対象となる可能性がある。

例えば、*SCHUFA I* 判決において、CJEU はドイツで使用されている**信用度スコアリング**システムを検討した。[113]この場合、コンピュータプログラムによって生成された「スコア」は、その人の支払約束を履行する能力に関する「確率値」であり、CJEU はこれを「プロファイリング」と認定した。より具体的には、このシステムは「ある人物のある特徴に基づき、ローンの返済など、ある人物の将来の行動の確率（「スコア」）に関する予言」を設定するものであった。

スコアの設定（「スコアリング」）は、ある人物を、ある行動をとった同等の特徴を持つ他の人物のグループに割り当てることで、同様の行動を予測できるという仮定に基づいている。[114]CJEU によれば、この活動は GDPR 第 4 条 4 項の意味における「プロファイリング」の定義を満たしている。[115]また、このようなプロファイリングは、AI 法第 5 条(1)(c)の意味における個人的特徴に基づく人物の評価に該当すると考えられ、AI システムを用いて行われる場合、同条項が適用される他の条件が満たされていれば禁止される。

b) 一定期間にわたって

[108]Article 29 Working Party, *Guidelines on Automated individual decision making and Profiling for purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7.

[109]GDPR 第 4 条 4 項および第 22 条、LED 第 11 条を参照のこと。第 29 条作業部会「規則 *2016/679* の目的のための自動化された個人の意思決定とプロファイリングに関するガイドライン」（WP251rev.01、2018.2.6、p.7）も参照のこと。

[110]特に、「プロファイリング」に言及している AI 法第 5 条(1)（d）の個別犯罪リスク予測の禁止や、AI 法第 5 条(1)（f）および（g）の感情認識およびバイオメトリクス分類は、特定の場合において禁止されている。

[111]AI 法 3 条 52 項には「プロファイリング」の定義があり、GDPR4 条 4 項の定義と相互参照している。

[112]Article 29 Working Party, *Guidelines on Automated individual decision making and Profiling for purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7.

[113]2023 年 12 月 7 日の司法裁判所の判決、*SCHUFA Holding (Scoring)*, C-634/21, EU:C:2023:957（以下、「*SCHUFA I* 判決」と呼ぶ）、パラグラフ 47 など。

[114]同 14 項（強調）。

[115]同 47 項

(155) AI 法第 5 条(1)(c)の禁止事項は、評価や分類が「**一定の期間」にわたる**データに基づいていることを求めている。このことは、アセスメントが、極めて特定の個別的な状況からのデータや行動による、一回限りの、あるいは一挙に行われる評価や等級付けに限定されるべきではないことを示唆している。同時に、禁止事項の範囲を迂回することを避けるため、この条件は、ケースのすべての状況を考慮して評価されることが重要である。

例えば、移民・亡命認可機関が難民キャンプで、カメラやモーションセンサーを含む様々な監視インフラを基盤に、部分的に自動化された監視システムを導入している。分析されたデータが一定期間に及び、例えば特定の個人（移民など）を評価し、彼らが逃亡しようとするリスクがあるかどうかを確認する場合、これは「一定期間に及ぶ」と認定され、他のすべての条件が満たされていれば、AI 法第 5 条(1)(c)の禁止が適用される可能性がある。

c) 社会的行動や、既知、推測、予測される個人的または性格的特徴に基づく

(156) AI 法第 5 条(1)(c)で禁止されている「評価」と「分類」は、(i) 個人または集団の社会的行動、(ii) 既知、推論、予測される個人およびパーソナ ルの特性、またはその両方、のいずれかに関する AI を活用したデータの処理（多くの場合、広範なもの基 づいたものでなければならない。データは、個人から直接プロバイダに提供されることもあれば、監視を通じて、サードパーティから、あるいは他の情報からの推論を通じて間接的に収集されることもある。

(157) 最初のシナリオに関して、「**社会的行動**」とは、一般的に行動、振る舞い、習慣、社会内での相互作用などを含む広範な用語であり、通常、複数の情報源（[116] ）からの行動に関連するデータポイントを対象とする。これには、文化的イベントへの参加、ボランティア活動など、社会的・私的な文脈における個人および個人の集団の行動だけでなく、例えば債務の支払い、特定のサービスを利用する際の行動など、ビジネスの文脈における社会的行動、さらには公的・私的事業体、政府、警察、法律との関係（例えば、人が交通ルールを守るかどうか）も含まれる。複数の文脈やデータポイントからの社会行動データは、同じ事業体によって集中収集されることもあるが、多くの場合、分散的に収集され、異なる情報源から組み合わされる。

(158) 第二のシナリオは、採点が**個人的または性格的特徴に基づいて**行われる場合であり、特定の社会的行動の側面が含まれる場合もあれば、含まれない場合もある。 個人的特性」には、例えば、性別、性的指向または性的特性、性別、性自認、人種、民族、家族状況、住所、収入、世帯構成員、職業、雇用その他の法的地位、職場での実績、経済状況、経済的流動性、健康、個人的嗜好、興味、信頼性、行動、場所または移動、借金のレベル、車のタイプなど、個人に関連する様々な情報が含まれる。[117]パーソナリティ特性」は、原則として個人的特性と同義に解釈されるべきであるが、パーソナリティと しての個人の具体的なプロフィールを作成することを意味する場合もある。パーソナリティ特性はまた、多くの要因に基づき、個人自身、他の人物、または AI システムによって生成される判断

[116]AI 法前文 31 を参照のこと。

[117]そのような特徴の例をいくつか挙げた AI 法前文 42 を参照のこと。

を意味する場合もある。AI 法では、パーソナリティ特性はパーソナリティ特性およびパーソナリティ特性と呼ばれることもある。[118] これらの概念は一貫して解釈されるべきである。

(159) 既知、推論、予測される」個人特性や性格特性は、区別が必要な異なるタイプの情報や個人データである。**既知の特性**」は、入力として AI システムにプロバイダされた情報に基づいており、ほとんどの場合、検証可能な情報である。対照的に、**「推論された特性」は**、他の情報から推論された情報に基づくものであり、推論は通常 AI システムによって行われる。**予測される特性」とは**、100%未満の精度でパターンに基づいて推定される特性である。推論」（または派生）データの概念は、連邦データ保護法におけるプロファイリングの文脈でも使用されているため、AI 法第 5 条(1)(c)で使用されているこれらの概念を解釈する際のヒントになるかもしれない。[119]これらの異なるタイプのデータの使用は、スコアリングの正確性と公平性に異なる影響を及ぼす可能性があるため、特に処理が不透明であったり、正確性の検証がより困難なデータポイントに依存している場合には、考慮される可能性がある。

4.2.2. 社会的得点は、無関係な社会的文脈における不利益もしくは不利な扱い、および／または社会的行 動の重大性に不当もしくは不釣り合いな扱いをもたらすものでなければならない。

***a)* 社会的得点と治療の因果関係**

(160) AI 法第 5 条(1)(c)の禁止が適用されるためには、AI システムによって、あるいは AI システムの支援を受けて作成された社会的スコアが、被評価者または被評価者の集団に**不利益または不利な取り扱いをもたらすもの**でなければならない。言い換えれば、待遇はスコアの結果であり、スコアは待遇の原因でなければならない。このようなもっともらしい因果関係は、有害な結果が 、まだ具体化していないが、AI システムがそのような不利な結果をもたらすことを意図しているか、もたらす可能性がある場合にも存在しうる。AI 法第 5 条(1)(c)の禁止行為が、このような AI システムの「上市」も対象としていることを考えると、この点は特に関連性が高い。

(161) AI 法第 5 条(1)(c)は、AI システムによって行われる評価や分類が、不利益または不利な取り扱いの唯一の原因であることを要求していない。従って、他の人間によるアセスメントの対象となったり、それと組み合わされたりする可能性のある、AI を活用した採点手法も対象となる。同時に、AI の出力は社会的スコアを生み出す上で十分に重要な役割を果たさなければならない。例えば、公的機関が人物の信頼性を評価するために AI システムを展開し、そのアウトプットを付加的な事実に関する人間のアセスメントと組み合わせる場合、この AI を活用した社会的スコアリングの慣行は、AI が生成的なスコアが最終的な決定において十分に重要な役割を果たす場合に限り、禁止事項の範囲に含まれる。

[118]AI 法第 5 条(1)（d）。

[119]Article 29 Working Party, *Guidelines on Automated individual decision making and Profiling for purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7 et seq.を参照のこと。

(162) スコアが、そのスコアを使用する組織とは異なる組織（複数可）によって作成されたものであっ ても、不利益または不利な取り扱いにつながる可能性がある[120] 。例えば、認可当局は、信用度やリスクアセスメントを専門とする別の会社が作成した自然人の信用度評価のスコアを取得することができる。

***b)* 関連性のない社会的文脈における有害または不利な扱い、および／または不当または不均衡な扱い**

(163) AI 法第 5 条(1)(c)の禁止が適用されるための最後の条件は、社会的スコアの使用が、不利益または不利な取り扱いをもたらす（またはもたらす可能性がある）ものでなければならないということである：

i. データが元々生成または収集された文脈とは無関係な社会的文脈において。

ii. 社会的行動やその重大性に不当または不釣り合いである。

(164) これらの条件は代替的なものであり、組み合わせて適用されることもある。多くの AI を活用した採点・評価手法がこれらの条件を満たさず、禁止事項の範囲外となる可能性があるため、少なくともいずれかが満たされているかどうかを評価するためには、ケースバイケースの分析が必要である。特に、AI を活用した採点・評価手法が特定の正当な評価目的のためのものであり、評価の目的に関連するとみなされるデータを特定し、不利益または不利な扱い（ ）が正当化され、社会的行動と比例することを保証する適用法（4.3.対象外参照）を遵守している場合は、この限りではないかもしれない。

(165) **「不利な扱い」とは、**採点の結果、その人や集団が他の人と比べて不利に扱われなければならないことを意味し、必ずしも特定の危害や損害を必要としない（例えば、不正の疑いがある場合に特別に追加検査の対象とするような採点方法の場合）。対照的に、「**不利益な**」待遇は、個人または集団がその待遇から特定の危害や不利益を被ることを必要とする。また、不利な扱いや不利益な扱いは、EU の無差別法で禁止されている識別的なものであったり、特定の個人や集団の排除を意味する場合もある[121] が、それは禁止が適用されるための必要条件ではない。したがって、AI 法第 5 条(1)(c)は、特定の保護対象集団（年齢、民族的・人種的出身、性別、宗教など）にのみ適用される EU 非差別法を超える不当な扱いをカバーすることができる。

シナリオ *1*：無関係な社会的文脈における不利益または不利な扱い

(166) AI 法第 5 条(1)(c)(i)に基づく最初のシナリオでは、点数に起因する不利益又は不利な取扱いは、データが当初生成又は収集された文脈とは**無関係な社会的文脈で**行われなければならない。このことは、社会的得点のために不利益または不利な扱いを受ける可能性があることを意味するだけでなく、その社会的行動または既知、推論もしくは予測される個人的もしくは人格的特性に関するデータが、得点が行われる社会的文脈とは無関係な社会的文脈で生成または収集されることも意味する。このような無関係な文脈から収集または生成されたデータは、その後、AI システムによって、評価や分

[120]この解釈は、CJEU が自動意思決定の文脈で、最終決定を行う事業体以外の事業体が作成した「スコア」（プロファイリングを構成する評価）が GDPR 第 22 条に基づく自動意思決定を構成しうるとした *SCHUFA I* 判決の CJEU 判決と一致している。*SCHUFA I* 判決の 42～51 項および 60～62 項を参照のこと。

[121]121 前文 31 AI 法。

類を目的とした明白な関連性なしに、または個人または個人グループの生成的監視につながる方法で、人物の採点に使用されなければならない。ほとんどの場合、このようなことは、本人の合理的な期待に反し、EU のデータ保護法および場合によっては、評価または分類に関連し必要であると考えられるデータの種類および情報源を規定するその他の適用規則に違反して行われる。この条件が満たされるかどうかは、評価の目的、データが収集され生成された背景を考慮して、ケースバイケースでアセスメントする必要がある。

第 5 条(1)(c)(i) AI 法で禁止されている、関連性のない社会的文脈における不利益または不利な扱いの例

- 国税当局は、国内の全納税者の納税申告書について AI 予測ツールを使用し、精査の対象となる納税申告書を選別している。AI ツールは、年収、資産（不動産、車など）、受取人の家族に関するデータなどの関連変数だけでなく、納税者の社会的習慣やインターネット接続などの無関係なデータも使用し、特定の個人を選別して検査を行う。

- ある社会福祉機関は、家計手当の受給者による不正の確率を推定するために AI システムを使用しているが、このシステムは、特定の国籍や民族出身の配偶者がいる、インターネットに接続している、ソーシャルプラットフォームでの行動、職場での成績など、不正のアセスメントに関連性や妥当性の明らかでない社会的背景から収集または推定された特徴に依存している。[122] これとは対照的に、公的機関は社会的給付が正しく配分されているかどうかを検証するという正当な目的を追求しているため、給付の配分に関連し、合法的に収集されたデータは、不正のリスクを判断するために使用することができる。

- ある公共労働機関では、AI システムを使って、面接と AI ベースのアセスメントに基づいて失業者に点数を付け、その個人が国の就労支援の恩恵を受けるべきかどうかを判断している。そのスコアは、年齢や学歴などの関連する個人データだけでなく、婚姻状況、慢性疾患の健康データ、依存症など、評価の目的とは明らかに関係のないデータや文脈から収集または推測される変数に基づいている。[123]

これらの容認できない採点方法は、EU 法および国内法に従って特定の目的のために個人を評価する合法的な方法とは区別される。

シナリオ 2：社会的行動と不釣り合いな、好ましくない、あるいは不利益な扱い

(167) AI 採点システムが禁止される可能性のある、AI 法第 5 条(1)(c)(ii)に基づくもう一つの代替シナリオは、採点の結果生じる処遇が不当であるか、社会的行動の重大性に不釣り合いである場合である。社会的行動の重大性と比較した社会的採点から生じる関係者の基本的権利への影響と干渉の

[122]似たような国の給付制度と社会的得点の比較については、D. Hadwick & S. Lan, 'Lessons to be learned from the Dutch childcare allowance scandal：オランダ、フランス、ドイツの税務当局によるアルゴリズム・ガバナンスの比較レビュー」（2021 年）World Tax Journal, Vol.13, Issue 4.Familiales (CNAF)を参照のこと。

[123]同様の制度はポーランドでも「失業者のプロファイリング」という制度で使われたが、違憲と判断され、放棄された。Szymielewicz, *Profiling the unemployed in Poland：Algorithmic Decision Making*, Fundacja Panoptykon, 2015, p. 18.

重大性によって、そのような扱いが追求される正当な目的に対して不釣り合いであるかどうかが、比例性の一般原則を考慮して決定されるべきである。これにはケース・バイ・ケースのアセスメントが必要であり、ケース・バイ・ケースのアセスメントでは、社会的行動の評価や不利益取扱いの比例性に関連する一般的な倫理的配慮や公正・社会正義の原則だけでなく、事案の関連するすべての状況を考慮すべきである。また、 合法的な目的を欠いているなど、「不当な」扱いである場合もある。このような潜在的な不利益扱いや不利な扱いを規制する具体的な規準や手続きを定めている部門別組合法または国内法も、このアセスメントの一環として関連する可能性がある。

第 5 条(1)(c)(ii) AI 法で禁止されている社会的行動と比較して不当または不釣り合いな扱いの例

- ある公的機関では、AI システムを利用して、親の精神的健康状態や失業などの規準に加え、複数の文脈から得られた親の社会的行動に関する情報に基づいて、リスクにさらされている子どもを早期に発見できるよう、家庭のプロファイリングを行なっている。その結果得られたスコアに基づき、家庭は検査の対象とされ、「リスクがある」と見なされた子どもは、医者の予約を時々欠席したり、交通違反の罰金を受けたりといった親の軽微な違反の場合も含め、家庭から引き離される。

- ある自治体では、AI システムを使って、さまざまな文脈における社会的行動に関する複数のデータに基づいて、住民の信頼度をスコア化する。信頼度が低い」と判断された住民の生成的スコアは、ブラックリスト化、すなわち公的給付の取り消し、その他の重大な懲罰的措置、管理・監視の強化に用いられる。アセスメントで考慮される要因の中には、ボランティア活動が不十分であることや、図書館に時間通りに本を返さない、収集日以外に路上にゴミを放置する、地方税の支払いが遅れるなどの軽微な不品行がある。

このような容認できない社会的得点の慣行は、特に、不利益または不利な扱いが正当化され、社会的行動と比例することをこれらの法律が保証している場合（4.3.適用範囲外参照）、欧州連合法および国内法を遵守し、合法的な特定の目的のために人を評価する合法的な慣行とは区別される。

(168) AI 法第 5 条(1)(c)(i)および(ii)に基づく両方の選択肢を同時に満たすこともできる。

AI 法第 5 条(1)(c)(i)および(ii)に基づく不当または不釣り合いな待遇の例

- ある税務当局は、低所得、二重国籍、社会的行動などの規準を使って、不正の疑いのある受給者をプロファイリングし、「故意／重大な過失」などのカテゴリーに振り分けることで、児童手当の不正を検知するために AI システムを使用している。リスクスコアに基づき、受給者のファイルが検査され、多くの場合、育児給付金は停止され、受給した給付金を返済するよう通知を受け、標準的な債権回収の手配を受ける資格はなくなる。このようなスコアリングは、多くの家庭に多額の負債を負わせ、個人および個人グループ[124] に対する不当、 差別的、不利益な扱いにつながり、多くの家庭を深刻な経済的苦境に追いやる。

[124]オランダの育児給付金スキャンダルの類似例については、*Belastingdienst treft 232 gezinnen met onevenredig harde actie*, 27.11.2019, (in Dutch)を参照のこと。オランダの裁判所は 2020 年、「Systeem Risico Indicatie（SyRi）」は違法であると判断した。こちらも参照のこと。パワープレイではなくフェアプレーである。Onevenredig harde anpak van 232 gezinnen met kinderopvangtoeslag, 2017, p. 32.

- ある認可当局は、学生寮の補助金プロセスにおける不正を管理するために AI システムを使用している。このシステムでは、不正リスクの識別要因として、受益者のインターネット接続状況、家族構成、教育レベルなどを指標としているが、これは関連性があるとは思えないし、正当化もされていない。

- ある政府が AI を使った総合的なシステムを導入し、社会的交流、オンライン活動、購買習慣、支払い時間の厳守など、生活のさまざまな側面における行動に基づいて市民を監視し、格付けする。スコアが低い人は、公共サービスへのアクセスが制限され、ローンの金利が高くなり、旅行やアパートの賃貸、就職さえも困難になる。このシステムは、個人の過剰な監視や、社会的スコアを決定するために使われる社会的行動とは無関係な文脈での不利益な扱い（例えば、就職の機会がソーシャルメディアの活動によって左右される）をもたらす一方、軽微な違反に対して過剰な罰則を課す（例えば、比較的軽微な違反に対して社会的・経済的不利益を被る）。

これらの容認できない社会的スコアリングの慣行は、これらの条件を満たさず、EU 法および国内法を遵守している合法的な特定の目的のために個人を評価する合法的な慣行と区別される場合があり、特に、これらの法律が、不利益または不利な取り扱いが正当化され、比例することを保証し、関連する社会的文脈からのデータが使用される場合である（4.3.対象外参照）。

(169) AI 法第 5 条(1)(c)に基づく禁止は、特定の個人または集団に賞罰や優遇措置が与えられる場合も対象となりうるが、これは他の個人に対する不利な扱いを意味するからである（例えば、雇用支援プログラム、住宅や再定住に対する（非）優先順位付けの場合）。

4.2.3. 公私の別を問わず、プロバイダが提供または使用する。

(170) すでに述べたように、AI 法第 5 条(1)(c)は、AI システムやスコアのプロバイダや利用者が公的か私的かを問わず、AI を利用した容認できない社会的採点行為を禁止している。公共部門における採点は、権力の不均衡や公共サービスへの依存により、人々に非常に重大な結果をもたらす可能性があるが、企業やその他の事業体による採点が増加している民間部門においても、同様の有害な結果が生じる可能性がある。

例えば、

- ある保険会社は、銀行から支出やその他の財務情報を収集する。これらの情報は、生命保険の候補者の適格性の決定とは無関係であり、そのような保険に支払われる保険料の価格を決定するために使用される。AI システムはこの情報を分析し、それに基づいて、特定の個人や顧客グループに対して契約を拒否するか、生命保険料を高く設定するかを推奨する。

- ある民間信用機関は、AI システムを使って人々の信用度を判断し、無関係な個人的特徴に基づいて、個人が住宅ローンを組むべきかどうかを決定する。

これらの容認できない社会的スコアリングの慣行は、これらの条件を満たさず、欧州連合法および国内法を遵守した、特定の正当な目的のために個人を評価する合法的な慣行とは区別される。

(171) 管轄の市場監視当局によるチェックの場合、AI システムの機能の透明性、データの種類とデータソースに関する情報を提供すること、評価または分類の目的で、スコアが使用される社会的背景に関連するデータのみが処理され、それらのデータが合法的に収集されたものであること、システムが意図されたとおりに機能していること、その結果生じる不利益または不利な取り扱いが正当化され、社会的行動に比例していることを保証することなどにより、AI 慣行が合法的かつ正当であることを証明することは、それぞれの責任の範囲内で、AI システムのプロバイダと展開者に課される。適用される法令を遵守し、システムに組み込まれ、運用中に適用される適切かつ比例的なセーフガードを遵守することは、適用禁止を回避するのに役立つと同時に、合法的かつ有益な目的（例えば、プロセスの有効性、サービスの質、安全性などを向上させるため）のために、人物の評価または分類のために AI システムを使用することを可能にする（4.3.適用範囲外を参照）。

(172) リスクの高い AI システム（例えば、必要不可欠な公共サービスや給付、クレジットスコアリングや信用度評価、移住などの分野）の要件を遵守することは、これらのリスクの高い分野で評価や分類の目的で使用される AI システムが、プロバイダや展開者がそれぞれの義務（リスクマネジメント、透明性、データガバナンス、基本的人権への影響評価、人的監視、モニタリングなど）を履行する際に考慮すべき、容認できない社会的スコアリングの慣行とならないことを保証することにもつながる。

4.3. 範囲外

(173) AI 法第 5 条(1)(c)の禁止は、自然人または人の集団の採点にのみ適用されるため、評価が個人の個人的または人格的特性や社会的行動に基づかない事業体の採点は原則として除外される。たとえ、場合によっては個人が間接的に採点に影響される可能性があるとしても（例えば、予算配分の場合、自治体の全市民）。しかし、事業体が、社会的行動または個人的もしくは人格的特徴に基づく自然人の集団の評価または分類を集計した総合得点に基づいて評価され、この得点がこれらの者（例えば、企業の全従業員、行動が評価された特定の学校の生徒）に直接影響を及ぼす場合、他のすべての条件が満たされれば、その慣行は AI 法第 5 条(1)(c)され、AI 法第 5 条(1)(c)のすべての条件を満たす個人を評価または分類するのでない限り、必ずしも AI が関与しない、個々の人間のスコアの単なる集計である。

(174) さらに、自然人の採点は常に禁止されているわけではなく、上記で分析したように、AI 法第 5 条(1)(c)の条件がすべて累積的に満たされる限定的な場合にのみ禁止される。特に AI 法前文 31 では、禁止は「EU 法および国内法に従って特定の目的のために行われる自然人の合法的な評価行為に影響を及ぼすべきではない」と言及している。例えば、信用スコアリングやリスクスコアリング、アンダーライティングは、金融・保険事業のサービスには不可欠な側面である。このような慣行は、他の合法的な慣行（すなわち、サービスの質と効率を改善するため、より効率的なクレーム処理を保証するため、特定の従業員評価、不正防止と検知、法執行、オンラインプラットフォームにおけるユーザーの行動のスコアリング）と同様に、合法的であり、AI 法およびその他の適用される EU 法および国内法に従って実施されるのであれば、それ自体は禁止されない。

(175) 言い換えれば、社会的スコアを生成する目的で、合法的な方法で、かつスコアに使用された個人データが収集された文脈と関連する特定の目的のために、個人を評価または分類する AI システムは、スコ

アの使用による不利益または不利な取り扱いが正当化され、社会的行動の重大性に比例するものである限り、禁止されない。[125]

(176) 信用スコアリング、マネーロンダリング防止などの分野別連邦法への準拠は、評価という特定の合法的目的に関連し必要であるとして使用できるデータの種類を特定し、 の取扱いが正当化され社会的行動に比例することを保証するものであるため、AI 慣行が AI 法第 5 条(1)(c)の禁止事項の範囲外であることを保証する可能性がある。

AI 法第 5 条(1)(c)の適用範囲外となる、EU 法および国内法に沿った正当な得点慣行の例：

- 債権者または信用情報機関が顧客の経済的信用力または債務残高を評価し、クレジットスコアを提供し、または信用力評価を決定するために使用する金融クレジットスコアリングシステムであって、顧客の収入および支出ならびにその他の財政的および経済的状況に基づいているものは、クレジットスコアリングの正当な目的に関連し、かつ、信用力評価における消費者の公正な取り扱いを確保するためにデータの種類および必要な保護措置を規定する消費者保護法[126] を遵守している場合には、AI 法第 5 条(1)(c)の適用範囲外となる。

- その評価が、サービスの文脈における取引行動やメタデータなどの関連データ、過去の履歴、その他不正のリスクを判断するために客観的に関連する情報源からの要因に基づくものであり、不利益な取り扱いが不正行為の結果として正当化され、かつ比例するものである場合、企業は金融詐欺について顧客を評価する正当な利益を有し、それらの慣行は禁止の影響を受けない。

- 保険契約者のリスクの高い運転行動に関連してテレマティクス・ベースの料金体系を提供する保険会社が使用する、運転者がスピード違反や安全運転を維持していないことを示すテレマティクス機器を通じて収集された情報は、保険料の引き上げが運転者のリスクの高い行動に比例している場合に限り、その運転行動によって引き起こされる事故のリスクが高いため、その保険契約者の保険料を引き上げるために使用することができる。

- AI システムの意図された正当な目的（例えば、患者を診断するために様々な情報源から収集された健康データや統合失調症データ）に関連し必要なデータの収集と処理は、特に、関連し必要なデータを処理するものであり、通常、特定の自然人に対する不当な不利益や不利な取り扱いを伴わないため、AI 法第 5 条(1)(c)の適用範囲外である。

- オンラインプラットフォームが、評価の文脈と目的に関連するデータに基づいて、そのサービスにおいて安全上の理由からユーザーをプロファイリングすることは、その評価がユーザーの不品行の重大性に不釣り合いな不利益な扱いをもたらさない場合、AI 法第 5 条(1)(c)の範囲外である。

- AI を利用したターゲティング商業広告は、関連データ（ユーザーの防御など）に基づき、消費者保護、データ保護、デジタルサービスに関する EU 法に沿って行われ、有害な 、またはユーザーの社会

[125]前文 31 AI 法。

[126]特に、消費者向け信用契約に関する 2023 年 10 月 18 日付指令（EU）2023/2225 および指令 2008/48/EC の廃止、ならびに 2020 年 5 月 29 日以降の融資組成およびモニタリングに関する欧州銀行認可のガイドライン（EBA/GL/2020/06）を参照のこと。

的行動の重大性に不釣り合いな不利な扱い（搾取的で不当な価格設定など）をもたらさない場合、対象外となる。

- 難民キャンプで収集されたデータ（行動コンプライアンスなど）を再定住や雇用の決定に使用する AI システムは、このデータが評価の目的に関連し、適用される EU 移民法に基づく手続きが履行され、その取り扱いが正当かつ比例的であることを保証するものであれば、禁止の影響を受けない。

- オンラインショッピングプラットフォームが、購入履歴が豊富で返品率の低いユーザーに対して、より迅速な返品申請プロセスや返品不要の払い戻しなどの特典を提供する AI を活用したスコアリングは、その特典が積極的な行動に報いるために正当かつ相応のものであり、他のユーザーは標準的な返品プロセスを引き続き利用できることから、AI 法第 5 条(1)(c)の対象外である。

- 個人の社会的行動に関するデータを複数のコンテクストから収集する警察その他の法執行機関による個人の AI 評価およびスコアリングは、それらのデータが刑事犯罪の予防、探知、訴追および処罰の特定の目的に関連し、不利益な取り扱いが実質的・手続的な連邦法および国内刑法・警察法に従って正当化され、かつ比例する場合には、AI 法第 5 条(1)(c)の適用範囲外となる。また、AI 法第 5 条(1)（d）の禁止事項についても考慮する必要がある。この禁止事項は、AI を活用したリスクアセスメントや犯罪を犯す可能性の予測について、プロファイリングや性格特性のアセスメントのみに基づいてはならないという、追加的かつ具体的な条件を課している（第 5 項参照）。

4.4. 他の連邦法との関係

(178) 特に、特定の評価目的に関連し必要であるとして使用できるデータの種類を厳格に規制する、より具体的な法律があるかどうか、正当かつ公正な取り扱いを確保するための、より具体的な規則や手続きがあるかどうかである。

(179) 企業対消費者関係において取引者として行動する私的当事者による AI を利用したソーシャル・スコアリング行為は、EU の消費者保護法、すなわち、不公正な企業対消費者商慣行に関する指令 2005/29/EC（「UCPD」）にも違反する可能性がある。UCPD は、商慣行が専門的な注意義務の要件に反し、製品に関する平均的な消費者またはグループの平均的なメンバーの経済行動を著しく歪めるか、または著しく歪める可能性がある場合、商慣行を禁止している（UCPD 第 5 条）。また、商慣習が消費者の取引上の意思決定に与える影響をケースバイケースで評価することを条件として、スコアリング行為が誤解を招くと判断される場合もある（UCPD 第 6～7 条）。

(180) 例えば、処理の法的根拠（合法性）、データ保護の原則（データの最小化と必要性、公平性、透明性など）、および関連する場合は、もっぱら自動化された個人の意思決定に関する規則を含むその他の義務に関して、公的機関であれ民間企業であれ、ソーシャル・スコアリングは EU のデータ保護法に違反する可能性もある。

(181) 評価や分類が、差別から保護される理由（年齢、宗教、人種や民族的出身、性別など）のいずれかに基づく場合、またはそれらの集団の差別を直接的または間接的にもたらす場合、そのような慣行も組合非差別法の対象となる。

(182) 消費者信用指令（EU）2023/2225[127] 。CCD 第 18 条 3 項は、信用度のアセスメントが、消費者の収入と支出、その他の財政的・経済的状況に関する適切かつ正確な情報に基づいて行われることを求めている。その情報には、収入やその他の返済原資の証拠、金融資産・負債に関する情報、その他の金融上の約束に関する情報などが含まれる。CCD は、特別なカテゴリーの個人データを情報に含めることや、ソーシャルネットワークから情報を取得することを明確に禁止している。欧州銀行監督機構（European Banking Authority's Guidelines on loan origination and monitoring）[128] 、信用力評価を目的とする関連情報をさらに規定している。これらの分野別法における特定の評価目的のためのデータの種類の特定は、ある慣行が AI 法第 5 条(1)(c)の禁止範囲に入るかどうかを判断する際に考慮すべき関連事項である。

(183) 同様に、アンチマネーロンダリングやテロ資金供与を目的とした人物の評価や分類に使用される AI システムも、これらの事項に関する EU の関連法規に準拠すべきである。

5. AI 法第 5 条(1)(d) -個々のリスクアセスメントと犯罪の予測

(184) AI 法第 5 条(1)d は、AI システムが、プロファイリングや人格的特徴・特性の評価のみに基づいて、自然人が犯罪を犯すリスクをアセスメントまたは予測することを禁じている。

(185) 同規定は最後のフレーズで、AI システムが、犯罪活動に直接関連する客観的かつ検証可能な事実にすでに基づいている、犯罪活動への人物の関与に関する人間のアセスメントを支援するために使用される場合は、禁止が適用されないことを示している。このような禁止範囲外の AI システム（ ）は、法執行当局、もしくはその代理、または法執行当局を支援する EU の機構、団体、事務所もしくは機関によって、プロファイリング、人格的特徴や特性、過去の犯罪行動の評価のみに基づかない、自然人の犯罪または再犯のリスクを評価するために使用されることを意図しており、「ハイリスク」AI システム（附属書 III、ポイント 6、レター（d）AI 法）として分類され、AI 法に基づくすべての関連要件および義務を遵守しなければならない。

5.1. 根拠と目的

(186) AI 法前文 42 は、AI 法第 5 条(1)(d)の禁止の背景と根拠を説明している。すなわち、自然人はその実際の行動に基づいて判断されるべきであり、AI が予測した行動に基づいて、そのプロファイリング、性格的特徴、特性のみに基づいて判断されるべきではないということである。

5.2. 禁止事項の主な概念と構成要素

***AI* 法第 *5* 条*(1)d* は次のように規定している。**

以下の AI 行為は禁止される：

[127]消費者向け信用契約に関する 2008 年 4 月 23 日付欧州議会および理事会指令 2008/48/EC、理事会指令 87/102/EEC の廃止、*OJ L 133,* 22/05/2008, p. 66-92.

[128]欧州銀行監督機構、2020 年 5 月 29 日以降のローンの組成とモニタリングに関するガイドライン、EBA/GL/2020/06。

d) 自然人が犯罪を犯すリスクを評価又は予測するために，自然人のプロファイリング又 は性格特性及び特徴の評価のみに基づき，自然人のリスクアセスメントを行うための AI システムを上市すること，この特定の目的のために稼働させること，又は使用すること。この禁止は，犯罪活動に直接関連する客観的かつ検証可能な事実に既に基づいている犯罪 活動への人の関与に関する人間の評価を支援するために使用される AI システムには適用されない；

(187) AI 法第 5 条(1)d の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「特定の目的での使用」、「使用」に該当しなければならない。

(ii) AI システムは、自然人が犯罪を犯すリスクを評価または予測するリスクアセスメントを行わなければならない。

(iii)リスクアセスメントまたは予測は、以下のいずれか、または両方のみに基づいていなければならない：

(a) 自然人のプロファイリング、

(b) 自然人の性格特性や特徴をアセスメントする。

(188) 禁止が適用されるためには、3 つの条件が同時に満たされなければならない。最初の条件、すなわち AI システムの上市、使用開始、使用については、2.3 節ですでに分析した。したがって、この禁止は、AI システムのプロバイダと展開者の両方に適用され、それぞれの責任の範囲内で、そのような AI システムを特定の目的のために上市、サービス開始、使用しない。禁止が適用されるための他の 2 つの条件については、以下で分析する。

5.2.1. 人が犯罪を犯すリスクをアセスメントし、その可能性を予測する。

(189) 個人が犯罪を犯すリスクを評価または予測するリスクアセスメントは、しばしば個人の「犯罪予 測」または「犯罪予測」と呼ばれる。犯罪予知」や「犯罪予測」の定義について、一般的に合意されたものはないが、[129] 、これらの用語は一般的に、犯罪学理論と組み合わせて、犯罪と闘い、犯罪を抑制し、犯罪を予防するための警察や法執行機関の戦略や行動に情報を提供するための基礎として、犯罪を予測するために使用される、多くの場合、大量の履歴データ（社会経済データだけでなく、警察記録なども含む）に適用される、さまざまな高度 AI 技術や分析手法を指す。[130]

[129]例えば、オランダの犯罪認知システム（CAS）やドイツとスイスのプレコブス（Precobs）など、EU 基本権庁のハンドブックに記載されているシステムを参照されたい（Handbook, 2018, p.138）。*Preventing unlawful profiling today and in the future: a guide*, Handbook, 2018, p.138.

[130]欧州刑事警察機構『AI and policing The benefits and challenges of artificial intelligence for law enforcement, An Observatory Report from the Europol』を参照のこと。

欧州刑事警察機構イノベーション・ラボ、2024 年 9 月 23 日。F. Yang, 'Predictive Policing' in *Oxford Research Encyclopedia, Criminology and Criminal Justice*, Oxford University Press, 2019 も参照のこと。

(190) 犯罪予測 AI システムは、過去のデータ内のパターンを識別し、犯罪発生の可能性と指標を関連付け、予測出力としてリスクスコアを生成する。例えば、このようなシステムは、警察のタスクフォースの計画、リスクの高い状況の監視、（再）犯罪の可能性が高いと予測される人物の取締りの実施に使用することができる。このようなシステムは、特に資源が乏しい法執行当局に機会をもたらし、効率を高め、犯罪を検知、抑止、予測するための積極的なアプローチを認可する。[131]しかし、他人の将来の行動を予測するために、犯した犯罪に関する過去のデータを利用することは、バイアスを永続させ、あるいは強化する可能性があり、重要な個人の状況がデータセットの一部でなかったり、特定の AI システムが作動するアルゴリズムで考慮されていなかったりする場合には、「見過ごされる」結果になりかねない。これはまた、法執行機関や司法制度全般に対する国民の信頼を損なう可能性もある 。[132]

(191) このようなリスクアセスメントや予測は、原則として、将来を見据えたものであり、将来の犯罪（まだ実行されていない）や、犯罪を実行するために行われた未遂や準備行為の場合を含め、現時点で実行されるリスクがあると評価されている犯罪に関するものである。□[133]□□れは、犯罪の予防や検知の間だけでなく、捜査、起訴、刑事罰の執行の間（司法当局が公判前勾留の賦課を決定する際など、再犯のリスクをアセスメントする場合を含む）、さらには刑事刑の服役後に社会へ再統合するための個人の計画の一部など、法執行活動のどの段階（[134] ）でも行うことができる。

(192) AI 法第 5 条(1)d の禁止は、犯罪予測やリスクアセスメントを違法とするものではない。それは、自然人が犯罪を犯すリスクを評価または予測するためのリスクアセスメントを行うための AI システムにのみ適用され、上記の第 3 の条件も満たされる。さらに、前述のとおり、AI 法第 5 条(1)d の最後のフレーズに含まれる明示的な除外に記載された状況では、禁止は適用されない。

5.2.2. 自然人のプロファイリング、または性格特性や特徴のアセスメントにのみ基づく。

(193) AI 法 5 条 1 項 d 号の禁止が適用されるための第三の条件は、自然人が犯罪を犯すリスクを評価または予測するためのリスクアセスメントが、a)その人のプロファイリング、または b)その人の人格的特徴や特性の評価のみに基づいていなければならないということである。

(194) AI 法第 5 条(1)d の禁止は、AI システムが一人の自然人のみ、あるいは複数の自然人の人格的特徴や特性を同時にプロファイリングまたはアセスメントするか否かに関係なく適用される。

a) 自然人のプロファイリング

[131]例えば、OxRec (Dutch Probation Office, 'Reclassering Nederland') Prediction of violent reoffending in prisoners and individuals on probation: a Dutch validation study (OxRec) - PMC (nih.gov) がある。

[132]例えば、EU 基本権庁（2022 年 12 月 8 日）Bias in algorithms - Artificial intelligence and discrimination｜欧州連合基本権庁を参照のこと。

[133]この点については、AI 法前文 42 の「犯罪を犯す可能性」と「実際の犯罪または潜在的な犯罪の発生」に言及しているが、これは現在形であって過去形ではない。

[134]一例として、児童の性的虐待と性的搾取および児童ポルノ対策に関する EU 指令 2011/93 の第 24 条 4 項がある。児童の性的虐待や性的搾取、児童ポルノとの闘いに関する EU 指令 2011/93 の第 24 条 4 項では、児童の性的虐待に関連する行為で刑事手続き中の者や有罪判決を受けた者に、再犯の危険性のアセスメントを受けるよう求めている。

(195) AI 法第 5 条(1)(c)とは異なり、第 5 条(1)(d)は「プロファイリング」という用語を明確に使用している。AI 法第 3 条第 52 項では、GDPR 第 4 条第 4 項（[135]）における定義を参照して、この用語を定義している。プロファイリングの概念には、「特定の個人的側面を評価する」という目的が中核的要素の一つとして含まれている。[136]AI 法第 5 条(1)(d)の文脈では、プロファイリングは、人が犯罪を犯すリスクを評価または予測する目的で行われる。

(196) この文脈では、いわゆる集団プロファイリング（[137]）の概念も関連する可能性がある。この概念は、例えば、犯罪の加害者のカテゴリー（テロリスト、ギャングなど）のような、他の人物が過去に犯した犯罪に関する過去のデータに基づいて構築された、所定のグループに対する記述的プロファイルの構築と適用を指す。これらのグループプロファイルは、他の人物が同様の犯罪を犯すリスクをアセスメントし、予測するために後で使用することができる。AI システムが予測を行い、そのような（グループ）プロファイルを特定の個人に適用する場合は常に、個人のプロファイリングに該当するため、AI 法第 5 条(1)（d）の禁止事項に該当する可能性がある。

***b)* 性格の特徴と特性のアセスメント**

(197) この禁止は、犯罪を犯すリスクを評価または予測するためのリスクアセスメントが、その人の人格的特徴や特性の評価のみに基づいている場合にも適用される。このようなアセスメントや予測は、しばしばプロファイリングの概念に含まれるが、GDPR 第 4 条第 4 項に定義されるプロファイリングが確立できない場合には、代替案とみなすこともできる。

(198) 4.2.1.c)節で述べたように、パーソナリティ特性および特徴は、特定の自然人に関連する特 徴の広範なカテゴリーを構成するものであり、一般的に合意された分類法は存在しない。前文 42 AI 法は、「国籍、出生地、居住地、子供の数、借金のレベル、車のタイプ」など、人が犯罪を犯すリスクを予測するた めにアセスメントされ得る性格特性及び特徴の例を示している。これは**あくまでも**例示であり、網羅的なリストではない。

(199) AI 法第 5 条(1)（d）は、同規定が対象とするリスクアセスメントが、人のプロファイリングまたは人格的特徴・特性の評価に「もっぱら」基づいている場合にのみ禁止されると規定している。AI 法第 42 条前文から明らかなように、「もっぱら」とは、プロファイリングにも、人格的特徴や特性の評価にも適用されることを意図している。

---

[135]AI 法第 5 条(1)d の禁止に関連する LED 第 3 条 4 項は、プロファイリングを同条と同じように定義している。

4(4)GDPR は、「自然人に関する特定の個人的側面を評価するため、特にその自然人の仕事上のパフォーマンス、経済状況、健康状態、個人的嗜好、関心、信頼性、行動、場所または移動に関する側面を分析または予測するために個人データを使用することからなるあらゆる形態のパーソナルデータの自動処理」としている。 同じ定義は、欧州連合の機構、団体、事務所および機関によるパーソナルデータの処理に関する規則（EU）2018/1725 の第 3 条（5）にも記載されている（OJ L 295, 21.11.2018, p. 39）。

[136]第 29 条データ保護作業部会、規則 *2016/679* の目的のための自動化された個人の意思決定とプロファイリングに関するガイドライン、WP251rev.01、2018.2.6、EDPB 承認、7 頁も参照のこと。Fundamental Right Agency, *Preventing unlawful profiling today and in the future: a guide,* Handbook, 2018, p.138 も参照のこと。

[137]集団プロファイリングについては、例えば、Fundamental Right Agency, *Preventing unlawful profiling today and in future: a guide,* Handbook, 2018, p. 21 を参照のこと。

(200) リスクアセスメントは、プロファイリングまたは性格特性や特徴の評価「のみ」に基づくものでなければならないという条件は、多くの状況において満たされない可能性がある。

(201) AI 法第 5 条(1)d 号の最後の文言から明らかなように、このような状況は、いずれにせよ、犯罪活動に直接関連する客観的かつ検証可能な事実にすでに基づいている犯罪活動への人の関与に関する人間のアセスメントを支援するために AI システムが使用される場合に生じる。前文 42 が明示しているように、この文脈では、特に、 、当該自然人に関して合理的な疑いがすでに存在している状況を考えるべきであるが、必ずしもそれだけに限定されるものではない。結局のところ、そのような場合、通常、関連する客観的かつ検証可能な事実に基づく人間によるアセスメントが行われているはずである。

(202) しかし、それ以外の状況もありうるため、常にケースバイケースでアセスメントする必要がある。一方では、「単独で」という用語の使用により、リスクアセスメントにおいて他の様々な要素が考慮される可能性が残されており、これはもはやプロファイリングや人格的特徴や特性の評価のみに基づくものではなくなっている。一方、禁止の迂回を回避し、その実効性を確保するためには、禁止が適用されないという結論を正当化できるような、現実的かつ実質的で意味のある要素が必要である。AI 法第 5 条(1)第 1 号（d）の禁止を、その最後のフレーズに含まれる除外規定とともに読めば、特に、事前に確立された客観的かつ検証可能な一定の事実の存在が、その結論を正当化しうることが示唆される。

例えば、

- ある法執行当局は、AI システムを使って、個人の年齢、国籍、住所、車のタイプ、配偶者の有無のみから、テロなどの犯罪行動を予測している。そのシステムにより、個人はその個人的特徴のみに基づいて、まだ犯していない犯罪を将来犯す可能性が高いとみなされる。このような制度は、AI 法 5 条 1 項 d 号で禁止されていると考えられる。

- 国税当局は、AI 予測ツールを使ってすべての納税者の納税申告書を確認し、潜在的な犯罪の可能性を予測して、さらなる調査が必要なケースを特定する。これは、二重国籍、出生地、子供の数などの性格的特徴や、不透明な変数、特に予測的であるため非客観的で検証が困難な推測情報をアセスメントに使用する、AI システムによって構築されたプロファイルに基づいてのみ行われる。このようなシステムは、特定の人物が犯罪行為に関与しているという合理的な疑いや、その犯罪行為に関連する客観的かつ検証可能な事実がないため、通常は AI 法 5 条 1 項 d の禁止事項に該当する。これはまた、AI 法 5 条 1 項 3 号で禁止されている社会的スコアリングの範囲に含まれる例である。

- ある警察署では、AI ベースのリスクアセスメントツールを使って、幼児や青少年が「将来、暴力犯罪や財産犯罪に巻き込まれる」リスクを評価している。このシステムは、他の人との関係や想定されるリスクレベルに基づいて子どもをアセスメントする。つまり、兄弟や友人など、リスクが高いと評価される人物とつながっているだけで、子どもは犯罪のリスクが高いと判断される可能性がある。両親のリスクレベルも、子どものリスク レベルに影響を与える可能性がある。リスクアセスメントの結果、警察はこうした子どもたちをシステムに「登録」し、追加検査で監視し、青少年「ケア」サービスに紹介することになる。このようなシステムも、AI 法第 5 条(1)d の禁止事項に該当する可能性が高い。

5.2.3. 犯罪行為に直結する客観的かつ検証可能な事実に基づく人間のアセスメントを支援するための AI システムの除外

(203) AI 法第 5 条(1)d は、その最後のフレーズにおいて、人の関与に関する人間のアセスメントを支援するために使用される AI システムには禁止は適用されないと規定している。

犯罪活動に直接関連する客観的かつ検証可能な事実にすでに基づいている。前述したように、この明示的な除外に記載された状況が、禁止が適用されない唯一の状況であるとは限らないが、この規定にその状況を明示的に含めることは、禁止の範囲を明確にし、その状況が問題となる場合には、いかなる場合にも禁止が適用されないことを明確にすることによって、法的確実性を提供する。

(204) システムが適用除外の範囲に含まれ、したがって禁止されていない場合、法執行当局が使用することが意図されている場合、または法執行当局に代わって使用されることが意図されている場合は、高リスク AI システム（附属書Ⅲ、AI 法 6(d)にいう）として分類され、したがって人間による監視を含む要件と保護措置の対象となる（AI 法第 14 条および第 26 条）。これらの要件には、人間による監視は、AI システムの能力と限界を正しく理解し、その出力を正しく解釈し、自動化のバイアスのリスクに対処できる、必要な能力、訓練、認可を受けた者に割り当てられなければならないことが含まれる。これらの人物は、明確な手順、トレーニング、AI システムの出力を有意義に評価するために必要な能力と認可を持つべきである。この具体的なケースにおいて、その人間によるアセスメントは、人が犯罪を犯すリスクに関する AI の予測やアセスメントが、犯罪行為に関連する客観的かつ検証可能な事実に基づいていることを保証するものでなければならない。また、これらの担当者は、悪影響やリスクを回避するために介入したり、AI システムが意図したとおりに作動しない場合にはその使用を停止したりすべきである。

(205) さらに、「人間の介入」という概念は、特に、航空旅客が重大犯罪に巻き込まれるリスクを予測する、もっぱら自動化された意思決定という文脈で、EU の判例の対象となってきた。この判例法は、AI 法第 5 条(1)d 号で使用されている「人間によるアセスメント」の概念の適用法にも関連する可能性がある。

Ligue des droits humains 事件（[138] ）において、EU は、航空旅行者の氏名記録（PNR）データを体系的に処理し、テロやその他の重大犯罪に関与する可能性を評価するための高度な AI システムの使用の合法性を検討した。 CJEU は、指令（EU）2016/681（「PNR 指令」）のルールを解釈し、自動化された処理のみに基づく不利な法的決定を禁止し、偽陽性を特定し、差別的でない結果を保証するために、非自動化された手段による陽性の一致に対して、**個別の人間による**アセスメントとレビューを要求した。 CJEU によれば、PNR データの自動処理結果の対象となる人間によるアセスメントは、テロ犯罪や重大犯罪に関与する可能性のある人物と一致するかどうかを評価する客観的な規準に基づくものでなければならず、自動処理の非差別性を保証するものでなければならない。

[138] 2022 年 6 月 21 日付司法裁判所判決、*Ligue des droits humains*, C-817/19, ECLI:EU:C:2022:491.

(206) 除外の内容に関しては、その中心的要素の一つは、AI システムが人間のアセスメントを 支援するために使用されることであり、禁止事項の対象となる状況で発生するような、AI システ ム自身がリスクアセスメントを行うことではない。しかし、この除外が適用されるためには、人間のアセスメントが、犯罪行為に直結する客観的かつ検証可能な事実に既に基づいていなければならない。

5.2.4. 民間主体の活動が適用範囲に入る可能性のある範囲

(207) 原則として AI 犯罪予測システムの主な展開先である法執行機関以外にも、場合によっては民間事業体の活動も AI 法第 5 条(1)d の禁止事項の対象となりうる。これは、その文言からして、この禁止が法執行機関のみに適用されるものではないという事実から導かれる。さらに、そうでなければ、この禁止は容易に迂回される可能性があり、その実効性が疑問視されることになる。

(208) そうである以上、特に、犯罪の予防、捜査、摘発もしくは訴追、または刑事罰の執行のために、公的権限および公権力の行使を法律により民間行為者に委託する場合には、この禁止が適用されるものと認可することができる[139] 。また、法執行当局に代わって行動し、個別の犯罪リスク予測を実施するよう、ケースバイケースで認可される場合もある。このような場合、適用される条件が満たされ、除外が適用されないのであれば、民間活 動者の活動も禁止の範囲に入る可能性がある。

> 例えば、高度な AI ベースの犯罪分析ソフトウェアを提供する民間企業が、法執行当局から、国家登録、銀行取引、コミュニケーションデータ、地理空間データなど、複数の情報源やデータベースから大量のデータを分析し、人身売買犯罪の潜在的犯罪者である個人（ ）のリスクを予測またはアセスメントするよう依頼される場合がある。第 5 条(1)(d)の規準がすべて満たされている場合、そのような使用例は禁止される可能性がある。

(209) さらに、犯罪を犯す人物のリスクをアセスメントまたは予測する民間事業体が、特定の犯罪を犯す人物のリスクをアセスメントまたは予測する法的義務を遵守するために客観的に必要である場合（例えば、マネーロンダリング対策やテロ資金供与の場合）には、この禁止が適用される可能性がある。

> 例えば、銀行機構は、EU のマネーロンダリング防止法（[140] ）に基づき、マネーロンダリング犯罪について顧客を選別し、プロファイリングする義務を負っている。銀行がその義務を果たすために AI システムを使用する場合、それは、容疑者として特定された人物がマネーロンダリング防止犯罪を犯す可能性が合理的に高いことを保証する客観的かつ検証可能な、同法に規定されたデータのみに基づいて行われるべきである。また、そのような評価の正確性と妥当性を確保するために、同法（[141] ）に従い、予測は人による評価と検証を受けなければならない。この法律を遵守することで、マネーロンダリング防止を目的とした個別犯罪予測 AI システムの利用は、AI 法第 5 条(1)（d）の禁止事項の範囲外となる。

[139] AI 法第 3 条 45 項の法執行当局の定義を参照のこと。

[140] 2024 年 5 月 31 日付のマネーロンダリング防止規則（EU）2024/1624。

[141] 141 規則（EU）2024/1624 第 20 条。

(210) しかし、禁止の文言から明らかなように、犯罪の実行に具体的かつ排他的に関連するリスクアセスメントに焦点が当てられていること、また、前文 42 で説明されている禁止の趣旨を考慮すると、民間事業体が、その通常の事業運営と安全のため、またはその財務上の利益を保護するため（例えば、財務上の不正を検出するため）に、特定の犯罪を犯すリスクを評価または予測する目的なしに、顧客をプロファイリングする場合、その民間事業体の活動は、前文 42 の禁止の範囲に含まれないと考えられる。顧客が特定の犯罪を犯すリスクをアセスメントまたは予測する目的なしに、民間事業体が通常の業務と安全のため、または財務上の利益を保護するため（例えば、財務上の不正を検知するため）に顧客をプロファイリングする場合、民間事業体の活動は AI 法第 5 条(1)（d）の禁止範囲に該当しないと考えられる。

(211) 言い換えれば、私的団体が法律により特定の法執行業務を委託されたり、法執行当局の代理を務めたり、上記のような特定の法的義務を負ったりしていない限り、私的団体の通常の事業過程において、自らの私的利益を保護する目的でリスクアセスメントを行うために AI システムを使用することは、そのリスクアセスメントが単に純粋に偶発的かつ二次的な状況として犯罪が行われるリスクに関連する可能性がある一方で、禁止事項の対象とはならない。

5.3. 範囲外

5.3.1. 位置情報ベース、地理空間情報ベース、場所ベースの犯罪予測

(212) 位置情報ベース、地理空間ベース、場所ベースの犯罪予測は、犯罪の場所や位置、あるいはそれらの地域で犯罪が起こる可能性に基づいている。原則的に、このような取り締まりは特定の個人のアセスメントを伴わない。したがって、これらは禁止事項の範囲外である。

位置情報ベース、地理空間的予測、場所ベースの犯罪予測の例

- AI ベースの予測的取り締まりシステムは、エリア別の過去の犯罪発生率や、ストリートマップなどのその他のサポート情報に基づいて、都市のさまざまなエリアで犯罪が発生する可能性のスコアを提供し、強盗やナイフ犯罪など、特定のタイプの犯罪のリスクが高いことを強調する。

- 税関当局は AI リスク分析ツールを使って、例えば既知の密売ルートに基づいて、麻薬や不正商品の所在地の可能性を予測する。

- ある警察署では、AI を活用したシステムによって銃声をリアルタイムで検知し、位置を特定している。このシステムは、都市部に設置された音響センサーによって銃声を識別し、その位置を三角測量することで、犯罪の検知と捜査に役立つ実用的なデータを警察官に提供している。

(213) しかし、位置情報に基づく犯罪予測システムと、人が犯罪を犯すリスクをアセスメントする個々の予測システムとをどのように区別するかは、必ずしも明らかではないかもしれない。ある AI システムが位置情報に基づく予測的取り締まりを行い、人物のプロファイリングにおける一側面として位置情報のリスクスコアを考慮する限りにおいて、そのシステムは人物ベースとみなされ、原則として AI 法第 5 条(1)（d）の対象となるはずであるが、他の理由で禁止の範囲から外れる可能性もある。

例えば、位置情報、地理空間情報、場所ベースの情報が、個人に関連する情報（例えば、ある人物の居住地）にリンクされており、AI システムが、犯罪が多発する居住地に基づくなど、当該個人のプロファイリングのみに基づいて、その人物が犯罪を犯す可能性が高いというリスクをアセスメントする場合、そのシステムは人物ベースとみなされるべきである。

5.3.2. 犯罪行為に関連する客観的かつ検証可能な事実に基づく人間のアセスメントを支援する AI システム

(214) AI 法第 5 条(1)(d)は、犯罪活動への自然人の関与に関する人間のアセスメントを支援する ために使用される AI システムには、禁止は適用されないと定めている。これは、 犯罪活動に直接関連する客観的かつ検証可能な事実にすでに基づいている。 このような場合、個々の犯罪リスクアセスメントや予測も、プロファイリングや個人的特徴のアセスメントのみに基づくものではないため、禁止されることはない。

このような理由で禁止の範囲外となる AI システムの例としては、以下のようなものがある：

- AI システムを実際の行動のプロファイリングや分類に使用すること。例えば、群衆の中で誰かが犯罪を準備中であり、犯罪を犯す可能性が高いと合理的に疑われる危険な行動などであり、AI の分類に対して人間による意味のあるアセスメントが行われる。この場合、AI の支援を受けて人間が行うリスクアセスメントは、個人的特徴やプロファイリングのみに基づくのではなく、行動を起こす前に人間が確認した、その人物の脅威的な犯罪行動に関連する客観的で検証可能な事実に基づく。

- 警察は武装強盗のリスクを調査しており、2 人の人物を疑っている。その疑いの根拠となる検証可能で客観的な事実がいくつか存在する。例えば、武器を購入するためのダークウェブのチャットグループへの参加や対話が検証可能である。地理空間予測または場所ベースの取り締まり情報と、容疑者が所有する車両の自動ナンバープレート登録（ANPR）情報を組み合わせた AI システムが、特定の犯罪行為に直結する検証可能かつ客観的な事実に基づいて、捜査における人間のアセスメントをサポートする。

- 囚人が早期釈放の恩恵を受けるべきかどうかのリスクをアセスメントする AI システムの使用。受刑者の AI プロファイルや性格特性や特徴のアセスメントは、過去の犯罪歴や更生に関連する実証された行動に関連する客観的かつ検証可能な事実の人間によるアセスメントを支援するものでしかない。

- 裁判官は、重大な犯罪で告発された人に対し、公判前勾留審理を行い、非親告罪の適用が可能かどうかをアセスメントする。その決定は、容疑者や被告人が勾留されなければ別の犯罪を犯す可能性や、逃亡したり捜査の適切な遂行を妨害したりする可能性など、公判前勾留を課す妥当性理由の有無のアセスメントに基づいて行われる。このプロセスを支援するため、裁判官は、類似の事件における個人の過去の犯罪歴や、年齢層、社会的行動、収入、雇用状況などの要素を含むデータに訓練された AI リスクアセスメントツールを使用する。

- AI システムは、非親告罪で服役している個人が釈放条件に違反したり、逃亡したりするリスクを、過去の犯罪行動や、釈放条件の遵守状況、精神アセスメントの結果、個人が利用している可能性の

ある他のコミュニティサービス（ ）からの推奨など、疑いを抱く根拠となる客観的事実に基づいて評価するために、人間の担当官のアセスメントを支援するために使用される。これらの情報に基づき、担当官は現状を維持するか、釈放条件を改定するかを決定する。

- 税関当局が使用する AI システムは、EU に入国する物品が国境で適用される法律（例えば、違法薬物の輸入禁止、輸出制裁違反、その他の違法行為を含む）を遵守していないリスクをアセスメントし、税関検査を実施すべき状況を特定するために使用される。AI システムは，物品及びそのサプライチェーンに関連して税関に提供された客観的かつ検証可能な情報（例えば，物品の性質及び価額，コンテナ番号，他の物品を隠すための輸送手段，特定の内容及び原産地の物品が同盟国への輸入又は同盟国からの輸出の要件に適合していることに関する予備知識）をアセスメントする。場合によっては、輸入事業者または輸出事業者が貨物の輸入に関連する不正行為に過去に関与していたこと、犯罪組織に所属していたこと、麻薬密売の犯罪歴があることに関する情報を処理することもある。なぜなら、自然人が不正商品の輸出入に関与する可能性の予測は、プロファイリングのみに基づくものではなく、商品および輸入事業者の犯罪活動への過去の関与に関連する客観的かつ検証可能な情報であり、税関管理またはリスク緩和措置が必要な状況か否かを判断するための人的審査の対象となるからである。

5.3.3. 事業体に関する犯罪予測やアセスメントに使用される AI システム

(215) AI 法第 5 条(1)(d)の禁止は、自然人の個別予測やリスクアセスメントにのみ適用されるため、通常、企業や非政府組織のような事業体をプロファイリングする犯罪予測システムは除外される。

例えば、

- 税務当局や税関当局は、AI システムを使って企業の取引や税務申告、通関データに関する大量のデータを分析し、企業が刑事犯罪を構成する税務や通関の不正行為を行うリスクをアセスメントしている。

- 税関当局を支援するために使用される AI システムは、不正商品を EU に送らないようにという認可を事業体に出すべき状況を特定するのに役立つ。

(216) 同時に、自然人が「個人事業者」または独立した専門家（弁護士など）として事業体を通じて行動する境界線上のケースもあり得る。このような場合、AI システムが特定の自然人をプロファイリングし、犯罪を犯すリスクをアセスメントまたは予測するため、すべての条件が満たされれば、AI 法第 5 条(1)d の禁止が適用される可能性がある。

5.3.4. 行政犯罪の個別予測に使用される AI システム

(217) AI 法第 5 条(1)d の禁止は、刑事犯罪の予見にのみ適用されるため、行政犯罪はその範囲から除外される。

例えば、行政当局が、軽微な犯罪（軽微な交通違反など）や税務、調達、支出プロセスにおける不正を犯す潜在的な犯罪者のリスクをアセスメントするために、行政調査の文脈で AI を使用する場合、

行政調査やチェックの結果、自然人が犯罪に関与する可能性のある情報が収集される可能性がある場合であっても、AI 法第 5 条(1)（d）の禁止の範囲には入らない。

(218) 犯罪の性質が行政犯罪か刑事犯罪かは、EU 法か国内法かによる。刑事犯罪」は EU 法の中で自律的な意味を持つ概念であり、加盟国間で一貫して解釈されるべきであるからである。CJEU は別の文脈において、加盟国による犯罪の分類はその点で決定的なものではないと結論付けている[142] 。犯罪の性質（犯罪か否か）を評価するために使用される関連規準は、CJEU および欧州人権裁判所（ECtHR）の関連判例に見出すことができる。[143]。

5.4. 他の連邦法との関係

(219) AI 法第 5 条(1)d 号の禁止事項と LED および GDPR との相互関係は、GDPR や LED などの連合データ保護法に基づくパーソナルデータの処理の合法性をアセスメントする際に関連する。特に、AI 法第 5 条(1)d 号は、法執行当局、その他の公的機関、およびその禁止範囲に含まれる民間事業体に対して、自然人のプロファイリングまたは人格的特徴および特性の評価のみに基づいて、自然人が犯罪を犯すリスクを評価または予測することの具体的禁止を課している。LED に関して、AI 法第 5 条(1)(d)は、LED 第 11 条第 3 項を損なうものではなく、（直接的または間接的な）差別をもたらすプロファイリングを禁止する。

(220) AI 法第 5 条(1)(d)の禁止と推定無罪に関する指令(EU)2016/343 の相互関係もまた関連している。というのも、 両法律は、指令の場合は直接的に、AI 法の場合は間接的に（その前文 42 を参照）、法律に従って有罪が証明されるま では無罪と推定されるという基本的権利に関係しているからである。[144]指令は人が刑事犯罪を犯したと疑われたり告発されたりした時点から適用される（[145] ）が、AI 法はより広い適用範囲を有し、特定の人に対する正式な刑事捜査が開始される前の予測や犯罪予防の段階ですでに適用される。また、そのような予測やリスクアセスメントが司法当局を含む管轄法執行当局ではなく、AI 法第 5 条(1)d)の範囲に属する私的行為者によって行われる場合にも適用される。

(221) AI 法第 5 条(1)d の禁止が適用されない場合であっても、適用される EU 法および国内法は、特にデータ保護法、刑事手続法、警察法、および個々の犯罪予測 AI システムの使用をさらに制限した

[142]例えば、2013 年 11 月 14 日付裁判所（大法廷）判決、Marián Baláž に対する違約金執行に関する訴訟、Case C-60/12, ECLI identifier : ECLI:EU:C:2013:733.

[143]CJEU の判例法によれば、いわゆる「エンゲル規準」に照らして、刑事罰でない刑罰が「刑事罰」とみなされるかどうかを判断するのは各国の裁判所である：ECtHR、1976 年 6 月 8 日判決、Engel and Others v. the Netherlands, Application nos. 5100/71, 5101/71, 5102/71, 5354/72 and 5370/72, CE:ECHR:1976:0608JUD000510071, paragraph 82.もともと欧州人権裁判所（ECtHR）によって開発され、その後 CJEU によって承認されたこれらの規準は、代替的なものであり、累積的なものではない。刑罰が犯罪的性質を有するか否かを検討する場合、管轄の国内裁判所は、(1) 国内法上の関連規定の分類、(2) 犯罪の性質、および (3) 刑罰の重大性を評価すべきである。犯罪の性質を評価する際に考慮される点としては、特に、訴訟手続が法的強制力を有する公的機関によって開始されているかどうか、その法的規則が懲罰的または抑止的な目的を有しているかどうか、その法的規則が刑法によって通常保護される社会の一般的利益を保護しようとしているかどうか、刑罰の賦課が有罪の認定に依存しているかどうかなどが挙げられる。刑罰の重さについては、国内法に定められている潜在的な最高刑が参考になる。これらの規準は代替的なものであり、必ずしも累積的なものではない。欧州人権裁判所、欧州人権条約第 6 条に関する手引き、公正な裁判を受ける権利（刑事）、2024 年 2 月 29 日更新を参照のこと。CJEU、2012 年 6 月 5 日判決、ボンダ事件 C-489/10、EU:C:2012:319、パラグラフ 37ff.、CJEU、2013 年 2 月 26 日判決、オーケルベルグ・フランソン事件 C-617/10、EU:C:2013:105、パラグラフ 35 も参照のこと。

[144]推定無罪は、EU 基本権憲章第 48 条に明記されている基本的権利である。

[145]CJEU が規定しているように、指令が適用されるためには、この人物が管轄当局によって容疑者／被疑者であることを知らされている必要はない。

り追加条件を課す可能性のあるセーフガードを含め、引き続き完全に適用されることを強調することが重要である。

6. AI 法第 5 条(1)(e) - 顔画像の非標的スクレイピング

(222) AI 法第 5 条(1)e 号は、インターネットや CCTV 映像から顔画像を無制限にスクレイピングして顔認識データベースを作成または拡張する AI システムの上市、特定目的での使用、使用を禁止している。

6.1. 根拠と目的

(223) インターネットや CCTV 映像から顔画像を無制限にかき集めることは、個人のプライバシー権やデータ保護権を著しく侵害し、匿名性を保つ権利を否定する。したがって、前文 43 の AI 法は、「大量監視の感覚」と「プライバシーの権利を含む基本的権利の重大な侵害」のリスクに基づいて、AI 法第 5 条(1)（e）に定める禁止を正当化している。

6.2. 禁止事項の主な概念と構成要素

> ***AI 法第 5 条(1)e 号は次のようにプロバイダを規定している。***
>
> 以下の AI 行為は禁止される：
>
> (e) インターネットや CCTV 映像から顔画像を無制限にスクレイピングし、顔認識データベースを作成または拡張する AI システムの上市、この特定の目的のためのサービス開始、または使用；

(224) AI 法第 5 条(1)e 号の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「特定の目的での使用」、「使用」に該当しなければならない；

(ii) 顔認識データベースを作成または拡張するためである；

(iii) データベースを作成する手段は、AI ツールによる非ターゲットのスクレイピングである。

(iv) 画像のソースはインターネットか CCTV 映像である。

(225) 禁止が適用されるためには、4 つの条件がすべて同時に満たされなければならない。AI システムの上市、使用開始、使用という最初の要素については、2.3 節ですでに分析した。したがって、この禁止は、AI システムのプロバイダと展開者の双方に適用され、それぞれがそれぞれの責任の範囲内で、そのような AI システムを上市、サービス開始、使用しないことになる。アンターゲッティング・スクレイピングの禁止に関する具体的な規準は、以下でさらに説明・分析される。この禁止は、インターネットや CCTV 映像から顔画像を非標的でスクレイピングするという特定の目的のために、上市または使用されるスクレイピングツールに適用される。つまり、この禁止は、顔認識用のデータベースを構築または拡張するためのスクレイピング・ツールには適用されず、対象を絞らないスクレイピングのためのツールにのみ適用される。

6.2.1. 顔認識データベース

(226) AI 法第 5 条(1)（e）の禁止事項は、顔認識データベースを作成または拡張するために使用される AI システムを対象としている。この文脈における「データベース」とは、コンピュータによる迅速な検索と取得のために特別に編成されたデータまたは情報の集合体を指すと理解されるべきである。顔認識データベースは、デジタル画像またはビデオフレームから顔を顔のデータベースと人間の照合し、データベース内の画像と比較し、両者が一致する可能性が高いかどうかを判断することができる。このような顔認識データベースは、一時的なものであっても、集中化されたものであっても、分散化されたものであってもよい。第 5 条(1)第 1 号（e）は、データベースの唯一の目的が顔認識に使用されることであることを要求しているのではなく、データベースが顔認識に使用されることができれば十分である。

6.2.2. 顔画像の非ターゲット・スクレイピングを通じて

(227) スクレイピング」とは、通常、ウェブクローラ、ボット、その他の手段を用いて、CCTV、ウェブサイト、ソーシャル・メディアを含むさまざまなソースから、 、自動的にデータやコンテンツを抽出することを指す。これらのツールは、データベースから情報を抽出し、その情報を別の目的に利用するためにプログラムされたソフトウェアである。

(228) 対象外」とは、「掃除機」のように動作し、可能な限り多くのデータや情報を吸収し、具体的かつ個別に意図されたデータ対象者を対象としない手法のことである。スクレイピングは、データやコンテンツを無差別に収穫する。従って、「対象を絞らない」という概念は、特定の個人やグループに焦点を絞らないことを意味する。robot.txt のようなインターネット・プロトコルのオプトアウトの尊重は、スクレイピングの非標的性には影響しない。

(229) スクレイピング・ツールが、特定の個人またはあらかじめ定義されたグループの人間の顔を含む画像または映像のみを収集するよう指示されている場合、スクレイピングは、例えば、特定の犯罪者を見つけるため、または被害者グループを特定するためなど、対象を絞ったものになる。このようなスクレイピングは、AI 法第 5 条(1)（e）の禁止事項の対象とはならない。

(230) 例えば、クローラーを使って、人身売買業者がソーシャル・メディア・チャンネルに投稿／広告する被害者の画像を拾い上げることで、あるクラスの被害者に焦点を絞った画像を収集することは、禁止事項の対象ではない。対象を絞らないスクレイピングは、禁止の迂回を許さない形で解釈されるべきである。インターネットや CCTV の映像を段階的にスクレイピングしてデータベースを作成し、その都度特定の個人グループやその他の規準を選択することは、最終結果が最初から対象を絞らないスクレイピングを追求するのと機能的に同じである場合、AI 法第 5 条(1)（e）の禁止事項に該当するはずである。

(231) システムが画像やビデオのターゲット検索と非ターゲット検索を組み合わせる場合、非ターゲットのスクレイピングは禁止される。

6.2.3. インターネットと CCTV 映像から

(232) AI 法第 5 条(1)（e）の禁止が適用されるためには、顔画像の情報源はインターネットか CCTV 映像のいずれかとなる。インターネットに関しては、ある人がソーシャルメディア上で自分の顔画像を公開

したからといって、その人が顔認識データベースにその画像が含まれることに同意したことにはならない。CCTV 映像から顔画像をスクレイピングする例としては、空港、道路、公園などの場所で運用されている監視カメラによって取得された画像が挙げられる。

例えば：

ある顔認識ソフトウェア会社が顔の写真を収集している。同社が保有する写真は、ソーシャルメディア（例：フェイスブック、ユーチューブ、

Twitter、Venmo）は、インターネットを検索し、人の顔を含む画像を検知する「自動画像スクレーパー」を使っている。これらの画像は、関連情報（画像のソース（URL）、地理的位置、場合によっては個人の名前など）とともに収集される。その後、顔の特徴が画像から抽出され、数学的表現に変換され、インデックス化と将来の比較のためにハッシュ化される。 ユーザーが個人の画像を AI システムにアップロードすると、AI システムはその画像がデータベース内の顔と一致するかどうかを判断する。アップロードされた画像は、スクレイピングされた画像と同じ数学的変換を受ける

(233) AI システムが人物の写真を受け取り、インターネット上でその顔に一致するものを検索する場合、つまり「画像検索エンジンのリバースエンジニアリング」は、標的を絞ったスクレイピングとみなされる。さらに、一致したものが「データベース」に表示されるかどうかは疑問である。

6.3. 範囲外

(234) AI 法第 5 条(1)（e）の禁止は、顔画像以外の生体データ（音声サンプルなど）の非標的スクレイピングには適用されない。 また、AI システムがスクレイピングに関与していない場合にも、この禁止は適用されない。人物の認識のために使用されない顔画像データベースも対象外である。例えば、AI モデルのトレーニングやテストの目的で使用される顔画像データベースなど、人物が特定されない場合は対象外である。

(235) AI 法第 5 条(1)e 号の禁止は、インターネットから大量の顔画像を収集し、架空の人物に関する新たな画像を生成する AI モデルを構築する AI システムには適用されない。そのような AI システムは、AI 法第 50 条の透明性要件に該当する可能性がある。

(236) AI 法第 5 条(1)（e）の禁止事項は、顔認識データベースの作成または拡張に使用される AI システムを対象としている。禁止の適用開始前に構築された既存の顔データベースで、AI を利用した非対象スクレイピングによってさらに拡張されないものに関しては、それらのデータベースとその使用は、適用される EU データ保護規則に準拠しなければならない。

(237) AI 法第 5 条(1)（e）の禁止は、顔認識データベースの構築や拡大を対象としている。バイオメトリクス識別の具体的な行為は、AI 法およびその他の関連する連合法の具体的な規則に従う。

6.4. 他の連邦法との関係

(238) EU のデータ保護法との関係では、顔認識データベースを構築または拡張するために、インターネットや CCTV の素材を無対象にスクレイピングすること、すなわちパーソナルデータの処理（データの収集とデータベースの使用）は違法であり、GDPR、EUDPR、LED に基づく法的根拠に依拠することはできない。

7. AI 法第 5 条(1)(f) - 行為感情認識

(239) AI 法第 5 条(1)f 号は、医療や安全上の理由による場合を除き、職場や教育機構の領域において、AI システムが自然人の感情を推測することを禁止している。禁止事項に該当しない感情認識システムは、AI 法附属書Ⅲの(1)(c)により、ハイリスクとみなされる。AI 法第 50 条第 3 項は、感情認識システムの使用について一定の透明性要件を定めている。

7.1. 根拠と目的

(240) 感情認識技術は急速に進化しており、人の感情を検知、収集、分析、分類、反応、対話、学習するためのさまざまな技術や処理操作を包括している。このような技術は「感情技術」とも呼ばれる。[146] 例えば、顧客行動の分析（[147]）、ターゲット広告やニューロマーケティング（[148]）などである。エンターテインメント業界では、パーソナライズされたレコメンデーションの提供や、映画に対する反応の予測などである；雇用の分野では、例えば採用プロセスに付随して、従業員の感情や退屈を監視するため、また「労働者をより幸せにする」ための幸福度アプリケーション[149]；法執行や公共の安全の分野では、例えば嘘発見器や大きなイベントでの感情スクリーニングのため、その他多くの目的のため。

(241) 感情認識は、その有効性や正確性がしばしば疑問視されている。[150]

特に、感情表現は文化や状況によって、また一個人であってもかなり異なるためである。このようなシステムの主な欠点として、信頼性の低さ、特異性の欠如、一般化可能性の低さが挙げられる」と説明している。さらに、感情認識は「差別的な結果をもたらし、関係者の権利と自由、特にプライバシー、人間の尊厳、思想の自由に対する侵害となりうる」と説明している。このことは、特に職場や教育訓練機構といった、労働者も学生も特に脆弱な立場にある非対称な関係において重要な役割を果たす。同

[146]感情を経済目的に利用することは、「エモーションミクス」とも呼ばれる。

[147]例えば、G. Mangano, A. Ferrari, C. Rafale, E. Vezzetti, F. Marcolin, 'Willingness of sharing facial data for emotion recognition: a case study in insurance market' in *AI & Society,* London, Springer, 2023 を参照のこと。

[148]N. Lee, A. J. Broderick, & L. Chamberlain, 'What is 'neuromarketing'?A discussion and agenda for future research', in *International Journal of Psychophysiology*, 63(2), 2007,199- 204,ニューロマーケティングを「市場やマーケティング取引に関連した人間の行動を分析し理解するための神経科学的手法の応用」と定義している(p.200)。

[149]E. Ackerman, & E. Strickland, 'Are you Ready for Workplace Brain Scanning?脳データを抽出・利用することで、労働者はより幸せになり、生産性も向上する、と支持者は言う」、*IEEE Spectrum*、2022 年 11 月 19 日、著者らは、「センサーは脳のさまざまな領域の電気的活動を検知し、その活動のパターンは、ストレス、集中、外部刺激への反応など、さまざまな感情や生理的反応と広く相関することができる」と説明している。

[150]例えば、J. Stanley, *Experts Say 'Emotion Recognition' lacks Scientific Foundation,* 18.7.2019, ACLU, refer to a study by L. Feldman Barrett e.a., 'Emotional Expressions Reconsidered : Challenges to Inferring Emotion From Human Facial Movements', *Psychological Science in the Public Interest*, 2019, pp.iii-90.

時に、安全や医療（健康治療や診断など）といった特定の利用場面における感情認識にも利点がある。[151]

7.2. 禁止事項の主な概念と構成要素

***AI* 法第 *5* 条*(1)*（*f*）にはプロバイダが規定されている：** 以下の AI 行為は禁止される： f) 職場や教育機構において、医療や安全上の理由から AI システムを使用することを意図している場合を除き、この特定の目的のために上市、使用開始、または自然人の感情を推測するために AI システムを使用すること。

(242) AI 法第 5 条(1)（f）の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「特定の目的での使用」、「使用」に該当しなければならない；

(ii) 感情を推測する AI システム[152] ；

(iii) 職場や教育訓練機構の領域で。

(iv) ただし、医療や安全上の理由による AI システムは禁止対象から除外される。

(243) 禁止が適用されるためには、4 つの条件がすべて同時に満たされなければならない。最初の要素、すなわち AI システムの上市、サービス開始、使用については、すでに 2.3 節で分析した。したがって、この禁止は、AI システムのプロバイダと展開者の双方に適用され、それぞれがそれぞれの責任の範囲内で、そのような AI システムを上市、サービス開始、使用しないことになる。禁止に関連するその他の条件については、以下でさらに説明・分析する。

7.2.1. 感情を推測する AI システム

a) 感情を推測する *AI* システムと感情認識システムの比較

(244) AI 法第 3 条第 39 項では、「感情認識システム」を「自然人の生体データに基づいて自然人の感情または意図を識別および推論することを目的とする AI システム」と定義している。AI 法第 5 条(1)（f）の禁止は、「感情認識システム」には言及しておらず、「自然人の感情を推測する AI システム」

[151]例えば、R. El Kaliouby and R. Picard and S. Baron-Cohen, 'Affective Computing and Autism', *Annals New York Academy of Sciences*、

2007, pp.228-248

[152]あるいは、（上市の際など）感情を推し量ることができる技術である。

のみに言及している。前文 44 はさらに、この禁止が「感情を識別または推論する」AI システムを対象としていることを明確にしている。

(245) 推論することは一般的に、前提条件として識別することを含むので、この禁止は、感情または意図を識別または推論する AI システムの両方を含むと理解すべきである。[153]一貫性の理由から、AI 法第 5 条(1)（f）の禁止は、他の感情認識システムに適用される規則（附属書 III 第 1 項（c）および AI 法第 50 条）と同様の範囲を有すると解釈し、個人の生体データに基づく推論に限定することも重要である。したがって、AI 法第 3 条第 39 項における感情認識システムの定義は、AI 法第 5 条(1)第 1 号（f）に関連していると考えるべきである

b) 感情や意図の識別と推論

(246) 識別」は、自然人の生体データ（例えば、声や顔の表情）の処理によって、感情認識システ ムにあらかじめプログラムされた感情と直接比較して識別することができる場合に行われる。推論」は、システム自身による分析その他の処理によって生成された情報を推論することによって行われる。この場合、感情に関する情報は、自然人について収集されたデータのみに基づいているわけではなく、感情を検出する方法をデータから学習する機械学習アプローチを含む、他のデータから推論される。[154]

c) 感情

(247) AI 法第 5 条(1)(f)の目的上、感情や意図の概念は広い意味で理解されるべきであり、制限的に解釈されるべきではない。AI 法第 18 条前文は、「幸福、悲しみ、怒り、驚き、嫌悪、困惑、興奮、恥、軽蔑、満足、娯楽などの感情」を列挙し、その詳細を説明している。これらの例は網羅的なものではない。

(248) この禁止は、態度に言及することで回避されるべきではなく、AI システムが生体データに基づいて、例えば人が怒っている態度を示していると判断する場合も含まれる。

(249) 前文 18 AI 法は、感情や意図には「痛みや疲労などの身体的状態（例えば、事故を防止する目的でプロのパイロットやドライバーの疲労状態を検知するために使用されるシステムを含む）」は含まれないことを明確にしている。さらに、感情認識システムには、「感情を識別または推論するために使用される場合を除き、容易に見て取れる表情、ジェスチャーまたは動作の単なる検出」は含まれないことを明確にしており、これは AI 法第 5 条(1)（f）にも適用されると理解すべきである。そのような表情とは、しかめっ面や笑顔などの基本的な表情、手や腕、頭の動きなどのジェスチャー、声の大きさやささやき声などの人の声の特徴などである。しかし、これらの容易にわかる表情やジェスチャーが、感情や意図の識別や推測に使われる場合は、禁止事項の対象となる。

例えば、

[153]AI 法前文 18 も参照のこと。

[154]AI 法前文 12 参照。したがって、推測されるデータは、データセットの相関関係やパターンの発見を目的とした、確率に基づく分析（ビッグデータ）プロセスの結果であることも多い。

- 人が笑っているという観察は、感情認識ではない。

- 病気かどうかの識別は感情認識ではない。

- テレビ局が、ニュース番組の司会者がカメラに向かって何回微笑んだかを記録する装置を使うのは、感情認識ではない。

- 人が幸せであると結論付けることが感情認識である。従業員が（例えば団体ジェスチャー、しかめっ面、笑顔の欠如から）顧客に対して不幸、悲しみ、怒っていると推測する AI システムは「感情認識」である。

- 声や団体のジェスチャーから、生徒が激怒して暴力を振るおうとしていることを推測するシステムは、「感情認識」である。

- AI 認識システムを使ってプロのパイロットやドライバーの疲労を推測し、事故回避のためにブレーキを踏むタイミングを示唆したり、注意を促したりすることは、「感情認識」ではない。感情認識には痛みや疲労といった身体的状態は含まれないからだ。

***d)* 生体データに基づいている。**

(250) AI 法第 3 条第 39 項の定義によれば、生体データに基づいて感情や意図を識別または推論する AI システムのみが感情認識システムに該当する。[155]

(251) 生体データを抽出できる個人特性は、身体的属性または行動的属性である。生理学的バイオメトリクスは、指紋、虹彩のパターン、顔の輪郭、手の静脈の形状など、人 の物理的、構造的、比較的静的な属性を用いる。モダリティの中には、本来は微視的であるが、それでも生物学的・化学的構造を示し、取得・識別が可能なものもある（DNA やにおいなど）[156] 。行動バイオメトリクスは、個人がタスクや一連のタスクを実行する際の動き、ジェスチャー、運動スキルの特徴的な特性を監視する。つまり、歩行（歩行分析）やキーボードへの指の接触（キーストローク）などの人間の動きが捕捉され、分析される。行動バイオメトリクスは、サイン、歩行、音声、キーストロークから、アイトラッキング、心拍[157] 、脳波（EEG）[158] 、心電図（ECG）[159] に至るまで、随意的および不随意的な反復運動と、それに関連する身体の特徴のリズミカルな タイミング／圧力を示す様々なモダリティを包含する。バイオメトリクス入力は、1 つのモダリティ（例えば、顔画像）または複数のモダリティ（例えば、脳波と組み合わされた

[155]AI 法第 3 条(34)：「生体データ」を「顔画像やダクティロスコピックデータなど、自然人の身体的、生理的、行動的特徴に関連する特定の技術的処理の結果として生じる個人データ」と定義している" AI 法前文 18 も参照。音声や会話からの感情推測について。

[156]生理学的および行動学的バイオメトリクス - バイオメトリクス機構

[157]生理学的および行動学的バイオメトリクス - バイオメトリクス機構

[158]EDPS、TechDispatch 1/2024 - Neurodata、3.6.2024 を参照。この中で、脳データと関連技術の利用が議論され、精神的プライバシーと完全性を含む新たな「ニューロライト」の提唱を含む法的意味合いも議論されている。S. O'Sullivan、H. Chneiweiss、A. A. による。

Pierucci and K. Rommelfanger, *Neurotechnologies and Human Rights Framework* : 報告書、OECD and CoE、9.11.2021、p.33、ニューロテクノロジーの現状と法的側面について論じている。

[159]Hasnul ら、2021、-Based ElectrocardiogramEmotion Recognition Systems and Their Applications in Healthcare を参照のこと

顔情報）に関連することができる。前文 18 では、顔の表情、手の動きなどのジェスチャー、または人の声の特徴を例として挙げている。

例えば、

- ある記事の文体や論調を定義するために、書かれた文章から感情を推測する AI システム（内容／感情分析）は、生体データに基づいているわけではないので、禁止の範囲には入らない。

- キーストローク（打ち方）、表情、体の姿勢や動きから感情を推測する AI システムは、生体データに基づくものであり、禁止事項に該当する。

(252) したがって、AI 法の生体データの定義は広範であり、感情認識、生体分類、その他の目的に使用されるあらゆる生体データが含まれる。[160]

7.2.2. 職場および教育機構への禁止事項の限定

(253) AI 法第 5 条(1)(f)の禁止は、「職場および教育機構の分野」での感情認識制度に限定されている。AI 法 44 条（ ）で明確にされているように、この制限は、職場や教育の場における権力の不均衡に対処するためのものである。

a) 「職場」

(254) 職場」という概念は広く解釈されるべきである。この概念は、例えば自営業の場合、自然人が雇用主や所属する組織から割り当てられた業務や責任に従事する、特定の物理的または仮想的な空間に関するものである。これには、仕事が行われるあらゆる環境が含まれ、屋内のオフィススペース、工場、倉庫から、店舗、スタジアム、博物館などの公共のアクセス可能なスペース、野外の現場や自動車、さらには一時的または移動可能な作業現場まで、仕事の性質によって大きく異なる可能性がある。これは、従業員、請負業者、研修生、ボランティアなどの地位とは無関係である。[161]AI 法第 5 条(1)（f）の「職場」という概念は、権限の不均衡があり、感情認識の押しつけ的性質が採用段階ですでに適用される可能性があるため、雇用、労働者管理および自営業へのアクセスの分野における AI システムの上市、稼働または使用を扱う AI 法の他の規定と一貫して、選考および採用プロセス中の候補者にも適用されると理解されるべきである。

例えば、

- コールセンターがウェブカメラや音声認識システムを使用して、怒りなどの従業員の感情を追跡することは禁止されている。[162]個人的なトレーニング目的でのみ展開される場合、その結果が人事責任者と

[160]AI 法では、生体データの定義に「一意の識別を可能にしまたは確認するもの」（生体データの機能的使用）という文言が含まれていないが、GDPR の生体データの定義にはこの要件が含まれている。生体データの GDPR の定義は、パーソナルデータの処理に関するデータ保護規則の下で適用される（そして、例えば GDPR 第 9 条 1 項および 9 条 2 項が適用される場合）。

[161]広範な解釈を展開している前文 56 など、職場における高リスクの AI システムに関連する説明も参照のこと。また、附属書 III の高リスクの AI システムのリストも参照のこと。

[162]Boyd et al., 2023, Automated Emotion Recognition in the Workplace.どのように提案された技術が可能性を明らかにするかの例：仕事の未来の

共有されず、トレーニングを受けた人のアセスメントや昇進などに影響を及ぼさないのであれば、感情認識システムは許可される。ただし、禁止事項が回避されず、感情認識システムの使用が職場関係に影響を及ぼさないことが条件となる。

- コールセンターが顧客の怒りや焦りなどの感情を追跡するために音声認識システムを使用することは、AI 法第 5 条(1)f 号で禁止されていない（例えば、従業員が特定の怒りっぽい顧客に対処するのを助けるため）。

- ハイブリッド・ビデオ通話の音声や画像から感情を識別・推測することで、ハイブリッド・ワークチームの感情的トーンを監視する AI システムは、通常、社会的認識の育成、感情的ダイナミクスの管理、紛争予防の目的に役立つはずだが、禁止されている。

- 採用プロセスで感情認識 AI システムを使用することは禁止されている。

- 試用期間中に感情認識 AI システムを使用することは禁止されている。

- スーパーマーケットが従業員の幸福感などの感情を追跡するためにカメラを使用することは禁止されている。

- スーパーマーケットや銀行が不審な客を検知するためにカメラを使用することは、例えば、誰かが強盗に入ろうとしていると判断するためであっても、従業員が追跡されておらず、十分な保護措置が講じられていることが保証されていれば、AI 法第 5 条(1)（f）で禁止されているわけではない。

***b)*「教育機関」**

(255) **教育機関に関する**言及は広範であり、公立・私立の両機関を含むものと理解されたい。児童・生徒の種類や年齢、特定の環境（オンライン、対面、ブレンデッド・モード[163] など）に関して制限はない。例えば、職業訓練校、すなわち生徒が手を使う技術を学ぶ学校（[164]）や継続的な訓練（[165]）を含め、あらゆるレベルの教育・訓練機構が AI 法第 5 条(1)（f）の禁止範囲に含まれる。教育機関は通常、関連する国の教育当局または同等の当局によって認可または認可されている。主な特徴は、教育機構が証明書を提供することである（それぞれ参加することが証明書取得の前提条件となる）。この禁止事項は、入学許可手続き中の候補者にも適用されると理解すべきである。

例えば、

[163]ブレンデッド・ラーニングとは、教育やトレーニングのプロセスにおいて、デジタルの融合を含め、複数のアプローチをとることであると理解される。

(オンライン学習を含む）および非デジタル学習ツール。

[164]例えば、欧州委員会の提案に付随する影響アセスメントでは、職業訓練機構による特定の AI 利用が、広範な基本的権利に強い干渉をもたらすとして、アセスメントの際などに言及されている：EU Commission, *Commission Staff Working Document.*影響アセスメント。附属書、*SWD(2021)84 final, Part2/2,* p. 43. I. Tuomi, *The , The use of Artificial Intelligence (AI) in education,* European Parliament, 2020, pp.9- 10 も参照のこと

[165]憲章第 14 条を参照のこと。

- 教育機関以外がオンラインで語学を学習するための感情認識を用いた AI ベースのアプリケーションは、AI 法 5 条 1 項 f 号で禁止されていない。これに対し、教育機関によって生徒がアプリケーションを使用することが義務付けられている場合、そのような感情認識システムの使用は禁止される。

- 教育機関が生徒をオンラインで試験する際に、AI ベースのアイトラッキング・ソフトウェアを使用して、目の凝視点や動き（注視点、例えば、無許可の教材が使用されているかどうかを検知するため）を追跡することは、システムが感情を識別したり推測したりするものではないため、禁止されていない。これに対し、感情の喚起や不安などの検知にも使用する場合は、禁止事項に該当する。

- 教育機関が生徒の興味や関心を推測するために感情認識 AI システムを使用することは禁止されている。これとは対照的に、ロールプレイの文脈における 学習目的（例えば、俳優や教師のトレーニングなど）のためにのみ展開される場合、その結果がトレーニングを受けている人の評価や認定に影響を与えないのであれば、感情認識システムは許可される。

- 教育機関が新入生の入学許可試験で感情認識 AI システムを使用することは禁止されている。

- 教育機関がオンライン講義中に、学生同士が携帯電話などを通じて会話している様子を撮影できる AI システムを使用することは、感情を推測するものではないため、禁止されていない。これに対し、感情の喚起、不安、関心などの検知にも利用する場合は、禁止事項に該当する。

- 教育機関が教師（職場）と生徒（教育）の両方に感情認識 AI システムを採用することは禁止されている。

7.2.3. 医療上および安全上の理由による例外

(256) AI 法第 5 条(1)（f）の禁止規定には、職場や教育機構で医療や安全のために使用される感情認識システム（治療用システムなど）に対する明確な例外規定がある。[166]高水準の基本的権利保護を確保するという AI 法の目的に照らせば、この例外は狭く解釈されるべきである。

(257) 特に、治療用途とは、CE マークを取得した医療機器の用途を意味すると理解すべきである。さらに、この例外は、ウェルビーイングの一般的側面を検知するための感情認識システムの使用には含まれない。職場におけるストレスレベルの一般的なモニタリングは、安全衛生の観点からは認められていない。例えば、職場や教育機構における燃え尽き症候群やうつ病の検知を目的とした AI システムは、この例外の対象とはならず、引き続き禁止される。

(258) この例外における安全上の理由という概念は、生命と健康の保護に関連してのみ適用されるものであり、他の利益、たとえば盗難や詐欺から財産を保護するために適用されるものではないと理解すべきである。

(259) この例外の狭い解釈から導かれるのは、医療上および安全上の理由による使用は、時間的、個人的、規模的な制限を含め、常に厳密に必要かつ適切なものに限定されるべきであり、十分な保護措

[166]166 前文 44 AI 法。

置を伴うべきであるということである。そのような保障措置には、例えば、特定の使用事例に関連する、書面による事前の動機づけられた専門家の意見を含めることができる。必要性は、医療・安全目的との関連において客観的な根拠に基づいてアセスメントされるべきであり、雇用主や教育機構の「ニーズ」に言及すべきではない。このアセスメント（ ）では、同じ目的を達成するためにより侵襲性の低い代替手段が存在するかどうかを問うべきである。

(260) 雇用者と教育者は、医療上および安全上の理由から、明確な必要性がある場合にのみ、感情認識システムを展開すべきである[167] 。この文脈で収集・処理されたデータは、他の目的に使用してはならない。職場での AI 管理ソフトウェアの使用が、労働者の健康と安全に。悪影響を及ぼす可能性があることが証明されていることを考えると、これは特に重要である例えば、ウェアラブルによる継続的な監視は、生産性に影響を及ぼすと同時に、作業ストレスを増大させる可能性がある 。[168]

(261) 前文 18 の AI 法では、感情認識システムの定義から痛みや疲労などの身体的状態を除外しているため、例えば、事故防止のためにプロのパイロットやドライバーの疲労状態を検知するために使用されるシステムなど、安全のために使用される多くの AI システムは、すでにこの定義に該当しないことになる。

(262) AI 法（[169] ）第 5 条(1)（f）の例外条件を満たす感情認識システムには、データ保護規則を含む他の法律が引き続き適用される。

(263) AI 法第 6 条第 2 項および附属書 III(1)(c)に基づき高リスクシステムに分類される感情認識システムは、AI 法第 3 章第 2 節の高リスク要件および AI 法第 50 条第 3 項の透明性義務を遵守する必要がある。

例えば、

感情認識は、自閉症の従業員や生徒を支援したり、目の不自由な人や耳の不自由な人のアクセシビリティを改善したりするために、医療上の理由で展開されることがある[170] 。このような使用は、AI 法第 5 条(1)（f）の医療上の理由による例外に該当する。

対照的に、学生や従業員の幸福度、モチベーションレベル、仕事や学習の満足度を評価するための感情認識は、「医学的理由による使用」には該当せず、禁止される。

雇用主は、測定されたストレスレベルに基づく不安の測定や、従業員の退屈度の測定のために、AI 対応機器やデジタルアシスタントを職場に展開することは禁止される。ただし、ストレスレベルの上昇や集中力の欠如が、例えば危険な機械の配備や危険な化学物質の取り扱いなど、特定の危険をもたらす

[167]EU 雇用法に従い、このような新技術を導入する場合、使用者は労働者またはその代表者と、国内手続きに則って協議しなければならない。このような手続き上の要件を尊重しなければ、AI 法を参照してこのようなシステムを導入することはできない。 データ保護法の観点からも同意が必要であり、これは引き続き適用される。

[168]AI 法と EU の労働安全衛生法枠組みとの相互関連 - Global Workplace Law & Policy (kluwerlawonline.com).

[169]2026 年 12 月以降、プラットフォーム作業における労働条件の改善に関する 2024 年 10 月 23 日付欧州議会および理事会指令（EU）2024/2831 が適用される。

[170]このシステムは、従業員や学生・生徒が同僚などの感情を理解するのに役立つだろう。

場合は例外である。後者の場合、雇用主は 、従業員の勤務成績のアセスメントなど他の目的にはデータを使用しないことができる。

7.3. より有利な加盟国法

(264) AI 法 2 条 11 項は、連合または加盟国は「使用者による AI システムの使用に関して、労働者の権利保護の観点から労働者により有利な法律、規制または行政規定」を維持または導入することができると規定している。また、労働者により有利な団体協約も許可または奨励される。

例えば、加盟国は、労働分野における感情認識システムの使用を医療目的に適用してはならないと定める法律を採択することができる。

7.4. 範囲外

(265) 前述したように、範囲外である

- AI システムは、生体データに基づいてではなく、感情や心情を推測する、
- AI システムは痛みや疲労などの身体状態を推測する。

(266) 職場や教育機構以外の領域で使用される感情認識システムは、AI 法第 5 条(1)f 号の禁止事項には該当しない。しかし、そのようなシステムはリスクの高い AI システムとみなされる。[171]同時に、このようなシステムは、AI 法第 5 条(1)(a)および(b)(有害な操作および搾取)、または他の連邦法によって、場合によっては禁止されることもある。このようなシステムには、EU データ保護法、消費者保護など、他のすべての適用法が引き続き適用される。

例えば、

顧客に対応するために商業的な文脈で使用される感情認識システムは、生体データに基づくか否かにかかわらず、AI 法第 5 条(1)（f）の禁止事項には該当しない。したがって、キーストロークに基づく、あるいは顧客の音声メッセージ（チャットメッセージ、バーチャル音声アシスタントの使用など）に基づく感情認識を可能にする AI システムのような例は、パーソナライズされたメッセージを表示するためのアプリケーションや、スマート環境（「インテリジェント看板」）を含む広告目的のオンラインマーケティングで使用されるが、禁止事項の対象とはならない。

とはいえ、このような行為は、AI 法（[172] ）第 5 条(1)（a）および（b）の有害な操作および搾取の禁止に含まれる可能性があり、これらの禁止が適用される条件がすべて満たされている場合に限られる。

a) 対象範囲外のその他のシステム

[171]AI 法第 6 条 2 項および付属書 III の 1 レターc）。

[172]こうした状況は、データや消費者保護など、他の規則でも禁止されている可能性がある。

(267) 「群衆統制」とは、（公共の）秩序やイベントの安全を維持するために、集団の行動を統制・監視することを指す。多くの場合、大規模な群衆イベント（サッカーやフットボールの試合、コンサートなど）や、空港や電車などの特定の場所に関連する。群衆統制システムは、例えばある場所の一般的な騒音やムードレベルを分析する場合、個々の人物の感情を推測することなく作動することができる。この場合、システムは（具体的な）自然人の感情を推論しないので、AI 法第 5 条(1)第 1 号（f）の適用範囲には入らない。

(268) しかし、そのような群衆制御システムが、例えば怒っている顔が多いかどうかなど、個人の感情を推測する場合もあるだろう。通常、このような AI システムは、職場や教育機構で使用されることはないため、AI 法第 5 条(1)（f）の禁止事項には該当しない。

(269) また、医療分野で使用されるシステム、例えば介護ロボットや、医療従事者が職場で検査中に感情認識システムを使用するシステム、緊急通報を分析する音声モニターなども対象外である。

(270) このようなシステムは、例えばサッカースタジアムや中央駅の警備員（攻撃的な行動を認識するためにこのようなシステムが使用される）、あるいは医療分野の従業員など、仕事の文脈でその場にいる人をスクリーニングすることが多い。このような場合、展開者は従業員の選別を避けるための安全策を採用しなければならない。しかし、このようなシステムが従業員の感情を推測することも完全に避けることはできない。システムの主目的は従業員の感情のアセスメントではないので、このようなシステムは禁止の範囲外であると考えるべきである。このようなシステムの展開者は、その使用によって従業員に悪影響が及ばないようにする責任を負っている。

8. AI 法第 5 条(1)(g) - 特定の「機微な」特徴に関するバイオメトリクス分類

(271) AI 法第 5 条(1)(g)は、人種、政治的意見、労働組合員、宗教的または哲学的信条、性生活または性 的指向を推測または推論するために、生体データに基づいて個人自然人を分類するバイオメトリッ ク分類システムを禁止している。この禁止は、例えば法執行目的で使用される可能性のある、連邦法または国内法に従って取得された生体データのラベリング、フィルタリング、分類は対象としていない。[173]。

8.1. 根拠と目的

(272) 「機微（センシティブ）」情報を含む様々な情報が、たとえ関係者が知らなくても、その人 を分類するために、バイオメトリクス情報から抽出、推論、推測される可能性がある。これは、例えば特定の人種であると見なされたためにサービスが拒否されるような、不公正で 差別的な扱いにつながる可能性がある。性的指向や政治的指向、人種といった側面に関連し、自然人を特定のグループやカテゴリーに割り当てることを目的とした AI ベースのバイオメトリクス分類システムは、人間の尊厳を侵害し、プライバシーや非差別といった他の基本的権利に重大なリスクをもたらす。したがって、これらは AI 法第 5 条(1)（g）によって禁止されている。

[173]173 前文 30 AI 法。

8.2. 禁止事項の主な概念と構成要素

AI 法第 5 条(1)g は次のようにプロバイダを規定している。
以下の AI 行為は禁止される： g) 人種、政治的意見、労働組合への加入、宗教的または哲学的信条、性生活または性的指向を推測または推論するために、生体データに基づいて個人自然人を分類する生体データ分類システムの上市、この特定の目的のための使用、または使用。この禁止は、合法的に取得された生体データセット（画像など）の生体データに基づくラベリングまたはフィルタリング、または法執行の分野における生体データの分類を対象としない

(273) AI 法第 5 条(1)（g）の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) その行為は、AI システムの「上市」、「特定の目的での使用」、「使用」に該当しなければならない；

(ii) このシステムは、バイオメトリック分類システムでなければならない；

(iii) 個人を分類しなければならない；

(iv) 生体データに基づいている；

(v) 人種、政治的意見、労働組合への加入、宗教的・哲学的信条、性生活、性的指向を推測・推論すること。

(274) 禁止が適用されるためには、5 つの条件がすべて同時に満たされなければならない。最初の条件、すなわち AI システムの上市、サービス開始、使用については、2.3 節で分析した。したがって、この禁止は、AI システムのプロバイダと展開者の双方に適用され、それぞれがそれぞれの責任の範囲内で、そのような AI システムを上市、使用開始、使用しないことになる。[174] 、禁止が適用されるその他の条件については、以下でさらに説明・分析する。

(275) この禁止は、法執行目的も含め、合法的に取得されたバイオメトリクス・データセットのラベリングやフィルタリングには適用されない。

8.2.1. バイオメトリクス分類システム

(276) (バイオメトリクス・システムによる個人の分類は、典型的には、個人の生体データが、ある事前に定義された特性を持つグループに属するかどうかを確定するプロセスである。個人の識別や身元確認ではなく、個人を特定のカテゴリーに割り当てることである。例えば、広告ディスプレイは、それを見ている個人

[174]AI システム」、「上市」、「この特定の目的のために使用されること」、または「使用」の基準については、上記を参照のこと。 [175]

の年齢や性別に応じて、異なる広告を表示することができる。[175]また、個人は、識別されることなく、識別する目的もなく、統計的な理由から単に分類されることもある。

(277) AI 法第 3 条(40)は、生体データ分類システムを、他の商業サービスに付随し、客観的な技術的理由により厳密に必要である場合を除き、その生体データに基づいて個人を特定のカテゴリーに 分類することを目的とする AI システムと定義している。7.2.1.d)で説明したように、「生体データ」は AI 法第 3 条(34)で定義されている。特に、生体データは、生体特徴に基づく行動特性から構成される。バイオメトリクス分類の範囲には、スカーフや十字架のような衣服やアクセサリー、 ソーシャルメディアの活動による分類は含まれない。

(278) バイオメトリクスによるカテゴライズは、身体的特徴（顔の特徴や形、肌の色な ど）のカテゴリーに依存する場合があり、それに基づいて、人が特定のカテゴリーに 割り当てられる。これらのカテゴ リーには、特別な「敏感な」性質のものや、人種など、EU の非差別的法律で保護される特 徴がある。しかし、バイオメトリクスによる分類は、DNA や、キーストロークの分析や人の歩 行のような行動的側面に基づくこともある[176] 。

(279) AI 法に基づくバイオメトリクス分類の定義の範囲から外れるためには、「他の商業サービスに付随し、客観的な技術的理由により厳密に必要である」という 2 つの条件が累積的に満たされなければならない。

(280) AI 法第 16 条によれば、純粋に付随的な機能とは、他の商用サービスと本質的にリンクしている機能であり、客観的な技術的理由により、主たるサービスなしにはその機能を使用できないことを意味し、その機能や特徴の統合は、AI 法の規則の適用を回避する手段ではない。

例えば、AI 法第 5 条(1)g 号では、以下のような AI の利用が認められている：

- オンラインマーケットプレイスで、消費者が自分自身で商品をプレビューできるようにするために使用される顔や身体の特徴を分類するフィルタは、商品の販売という主要なサービスに関連してのみ使用できるため、そのような付随的機能に該当する可能性がある。

- 顔や体の特徴を分類し、ユーザーが写真や動画を追加・修正できるようにする、オンライン・ソーシャルネットワーク・サービスに組み込まれたフィルターも、付随的な機能であると考えられる。

一方、禁止される用途の例としては、以下のようなものがある：

- ソーシャル・メディア・プラットフォーム上で活動する個人を、彼らがそのプラットフォーム上にアップロードした写真の生体データを分析することによって、想定される政治的指向に従って分類し、ターゲットとなる政治的メッセージを送る AI システム。このようなシステムは政治的広告に付随するものでしかないかも

[175]第 29 条作業部会、バイオメトリクス技術の発展に関する意見 *3/2012*、WP193、2012 年 4 月 27 日、6 頁を参照のこと。

[176]例えば、第 29 条作業部会、バイオメトリクス技術の発展に関する意見 *3/2012*、WP193、2012 年 4 月 27 日、16～17 頁を参照のこと。同グループはここで、「ソフトな認識」（p.17）、すなわち「人々の行動や特定のニーズの検知」に言及している。

しれないが、「客観的な技術的理由で厳密に必要」ではないため、バイオメトリクス分類の定義から除外する条件は満たされない。

- ソーシャル・メディア・プラットフォーム上で活動する個人を、そのプラットフォーム上で共有された写真の生体データを分析することによって、想定される性的指向に従って分類し、それに基づいてそれらの個人に広告を提供する AI システムは、AI 法の意味における生体情報分類に該当する。また、この場合、この「付随的サービス」には厳密な必要性がないため、禁止からの除外は適用されない。

8.2.2. 個人は生体データに基づいて個別に分類される。

(281) 自然人の分類のために生体データを使用することは、禁止が適用されるための不可欠な要素である（上記 8.2.1.および 7.2.1.d)項参照）。

(282) さらに、禁止が適用されるためには、自然人が「個々に」分類されなければならない。これがバイオメトリクス分類の目的でも結果でもない場合、禁止は適用されない。例えば、個人を見ずにグループ全体を分類する場合などである。

個々のカテゴリー分けの例としては、以下のようなものがある：

- 例えば、顔、身長、肌、目、髪の色などの身体的特徴（またはそれらの組み合わせ）に基づいて、例えば「年齢、性別、民族性」などの「属性推定」（人口統計をカウントする）を行う AI システム。

- AI システムは個人を分類し、特定の特徴（例えば、右目の下に傷跡がある）や右手にタトゥーがあるという理由に基づいて、その人を選別することができる。

これらの使用例は、個々のバイオメトリクス分類の例である。これらの例が AI 法第 5 条(1)（g）の禁止事項に該当するためには、同条項のすべての条件が満たされなければならない。

8.2.3. 人種、政治的意見、労働組合への加入、宗教的または哲学的信条、性生活、性的指向を推測または推論すること。

(283) AI 法第 5 条(1)(g)は、人種、政治的意見、労働組合員、宗教的または哲学的信条、性生活または性的指向といった、限定された機微な特徴を推測または推論することを目的とするバイオメトリクス分類システムのみを禁止している。

たとえば、AI 法第 5 条(1)（g）で禁止されているシステムには、以下のようなものがある：

- 個人の声から人種を推測できると主張するバイオメトリクス分類システム（これは、肌や目の色によって人を分類するシステムや、犯罪被害者の DNA を出自から分析するシステムとは異なる）。それらのシステムは禁止されていない）。

- というバイオメトリクス分類システムは、タトゥーや顔から個人の宗教的指向を推測できると主張している。

8.3. 範囲外

(284) AI 法第 5 条(1)(g)の禁止は、法執行の分野を含め、合法的に取得された生体データ（画像など）に基 づくラベリングやフィルタリングを行う AI システムを対象としていない。このことは、AI 法（[177] ）の前文 30 でさらに説明されている。

(285) 生体データセットのラベリングまたはフィルタリングは、データがすべての人口統計学的集団を等しく代表し、例えばある特定の集団を過剰に代表しないことを保証するために、生体分類システムによって正確に行われることがある。アルゴリズムの訓練に使用されるデータが特定の集団に対してバイアスされている場 合（すなわち、データの収集方法によって集団間にデータの系統的な差が存在する場合、 またはデータが歴史的に偏っている場合）、アルゴリズムはこのバイアスを再現する可能性があ り、その結果、個人または集団に対する違法な差別が生じる可能性がある。[178]このような理由から、差別を防ぐために、保護されるべき機密情報に基づくラベリングが、高品質なデータには必要かもしれない。AI 法は、リスクの高い AI システムに対して、AI 法の要件に適合したラベリング操作を要求することさえある。[179]したがって、このような生体データのラベリングやフィルタリングは、AI 法第 5 条(1)（g）の禁止事項から明確に除外されている。この禁止が適用されるのは、生体データが人種、政治的意見、労働組合員、宗教的または哲学的信条、性生活または性的指向を推測するために分類される場合のみである（ ）。

許容される表示やフィルタリングの例には、以下が含まれる：

- 生体データのラベリングにより、ある民族グループのメンバーが就職面接に招かれる確率が低くなるようなケースを回避する。これは、その特定のグループの成績が他のグループよりも悪い、つまり結果が悪いデータに基づいてアルゴリズムが「訓練」されたためである。[180]

- 皮膚や目の色によって患者を分類することは、医療診断、例えば癌の診断に重要かもしれない。

(286) AI 法第 5 条(1)f 号はまた、法執行の分野において合法的に取得されたデータセットのラベリングやフィルタリングには、同条項の禁止は適用されないと定めている[181] 。

例えば、児童への性的虐待が疑われるデータセットのラベリングとフィルタリングを可能にする AI システムを、法執行当局が使用することが対象となる。まず第一段階として、法執行機関は AI システムのサポートを利用して、画像からセンシティブなデータを検知し、再編集する。さらに、性別、年齢、目の色や髪の色などの生体データ、傷跡、マーキングなどに応じてフィルタリングやラベリングを行えば、被害者の特定や他の事件との関連付けに役立つ可能性がある。同様に、容疑者の特定に役立てるため、指の

[177]前文 30 AI 法：その禁止は、例えば法執行の分野で使用され得る、髪の色や目の色による画像の選別のような、生体データに従って、連邦法または国内法に沿って取得された生体データセットの合法的なラベリング、フィルタリング、分類をカバーすべきではない」。

[178]同上。

[179]例えば、AI 法第 10 条および第 17 条を参照のこと。

[180]FRA, # *BigData.Data supported decision making*, Luxemburg, 2018, 14, p. 5.

[181]法執行のためのバイオメトリクス分類システムの使用に関する AI 法は、TFEU 第 16 条に基づいている。AI 法の前文 3 も参照のこと。

長さ、特徴的なマークやタトゥーなど、特定の特徴に基づいて虐待者の手をフィルタリングし、ラベリングすることも認められている。

8.4. 他の連邦法との関係

(287) 生体データに基づいて GDPR 第 9 条(1)に基づき保護される機微な属性または特徴に従った生体認証分類に使用されることを意図した AI システムは、本規則で禁止されていない限り、AI 法[182]に基づき高リスク[183] に分類される。

(288) AI 法第 5 条(1)（g）は、GDPR、LED、EUDPR などの連合データ保護法に基づく合法的な個人データの処理の可能性をさらに制限する。特に、AI 法第 5 条(1)g 号は、人種、政治的意見、労働組合員、宗教的または哲学的信条、性生活または性的指向を推測するために、AI 法に定義された生体データに基づいて自然人を分類する可能性を排除する。ただし、法執行の分野を含む合法的に取得された生体データセットのラベリングまたはフィルタリングの例外は、前述（ ）のとおりである。さらに、AI 法第 5 条(1)（g）の禁止は、人種、民族的出身、性的指向、政治的意見、宗教的信条など、個人データの特別なカテゴリーに基づく識別的差別をもたらすあらゆる「プロファイリング」を明確に禁止している LED 第 11 条 3 項と一致している。

9. AI 法第 5 条(1)(h) - 法執行目的のリアルタイム遠隔バイオメトリクス識別（RBI）システム

(289) AI 法第 5 条(1)項(h)は、AI 法に網羅的に規定された限定的な例外を除き、公共のアクセス可能な空間におけるリアルタイム RBI システムの法執行目的での使用を禁止している。具体的には、AI 法第 5 条(1)(h)(i)～(iii)は、国内法によって認可され、AI 法第 5 条 2 項～(7)の条件と保護措置が満たされる場合、そのようなシステムの使用が許可される 3 つの状況を想定している。

(290) AI 法第 5 条(5)に従い、加盟国は、法執行のために公衆がアクセス可能な空間におけるリアルタイム RBI システムの使用が、自国の領域において許可されるかどうか、またどのような状況において許可されるかを自由に決定することができる。このような使用を許可し規制する国内法がない場合、法執行当局およびその代理を務める事業体は、法執行目的でこのようなシステムを展開することはできない。したがって、AI 法の関連要件に準拠する国内法が存在することが、このような利用の前提条件となる。

(291) AI 法第 5 条(1)(h)は、公共のアクセス可能な空間におけるリアルタイム RBI システムの使用を法執行目的でのみ禁止しているため、この条項が関係するのはそのようなシステムの展開者のみである。このようなシステムの上市および使用開始は、他の RBI システムの使用と同様、禁止されていないが、AI 法（[184] ）附属書 III の第 6 条(2)項および a)項に従った高リスクの AI システムに関する規則に従う必要がある。加盟国が、AI 法第 5 条(1)(h)に列挙された 3 つの目的のいずれかについて、法

[182]前文 54 および附属書 III 第 1 レターb）。AI 法。

[183]前文 54 および附属書 III 第 1 レターb）AI 法。

[184]さらに、法執行目的での RBI システムの遡及的使用に適用される特定の規則がある（AI 法第 26 条 10 項）。

執行のために一般にアクセス可能な空間におけるリアルタイム RBI システムの使用を許可する場合、高リスク AI システムに関する規則もその使用に適用される。

(292) 最後に、法執行を目的とした RBI システムの遡及的使用には、特定の規則が適用される。このような非リアルタイムの利用は禁止されていないが、リスクの高い AI システムの展開には追加的な保護措置が適用される（AI 法第 26 条 10 項）。

9.1. 根拠と目的

(293) 前文 32 AI 法は、法執行を目的とした公共のアクセス可能な空間におけるリアルタイムの RBI システムの、関係者の権利と自由に対する侵入的性質を認めている。それは、人口の大部分の私生活に影響を及ぼし、常時監視されている感覚を喚起し、間接的に 集会の自由およびその他の基本的権利の行使を妨げる可能性があるという程度である。自然人の遠隔バイオメトリクス識別を目的とする AI システムの技術的な不正確さは、バイアスのかかった結 果をもたらし、差別的効果を伴う可能性がある。そのような偏った結果や差別的効果の可能性は、年齢、民族、人種、性別、障害に特に関連する。さらに、リアルタイムで作動するこのようなシステムの使用に関連して、影響の即時性、さらなるチェックや訂正の機会の限定は、法執行活動との関連において、あるいは法執行活動によって影響を受ける関係者の権利と自由に対するリスクを高めている。

(294) しかし、そのようなシステムの使用が実質的な公共の利益を達成するために厳密に必要であり、そのような使用が起こりうる状況が網羅的に列挙され、狭く定義されている場合、その使用は基本的権利に対するリスクを上回る（AI 法 33 条）。このようなシステムが「責任ある適切な方法」で使用されることを保証するため、その使用は、AI 法第 5 条 2 項～7 項のセーフガードと具体的な義務および要件の対象となる。

9.2. 禁止事項の主な概念と構成要素

AI 法第 5 条(1)h 号

以下の AI 行為は禁止される：

h)ただし、そのような使用が以下の目的のために厳密に必要である場合を除く：

i) 拉致、人身売買の特定の被害者を探すこと。

ii) 自然人の生命若しくは身体の安全に対する具体的、実質的かつ差し迫った脅威、又はテロ攻撃の真正かつ現存する若しくは真正かつ予見可能な脅威の防止；iii) 附属書 II に規定される犯罪であって、当該加盟国において拘禁刑又は最高 4 年の拘禁令により処罰される犯罪について、犯罪捜査若しくは訴追を行い、又は刑事罰を執行する目的で、犯罪を犯したと疑われる者の所在を特定し、又はその身元を確認すること。

第 1 号の(h)は、法執行以外の目的での生体データの処理に関する規則(EU)2016/679 の第 9 条を損なうものではない。

(295) AI 法第 5 条(1)（h）の禁止が適用されるためには、いくつかの累積的条件が満たされなければならない：

(i) AI システムは RBI システムでなければならない；

(ii) 活動は、そのシステムの「使用」からなる

(iii)「リアルタイム」で、

(iv)公共のアクセス可能な空間、(v)法執行目的。

(296) 番目の条件、すなわち AI システムの「使用」については、本ガイドラインの 2.3.項ですでに分析済みである。上記の他の条件については、以下でさらに説明し、分析する。

9.2.1. 遠隔生体認証の概念

(297) 生体認識技術は、測定可能な身体的特徴（目の距離や大きさ、鼻の長さなど）または行動的 特徴（歩行や声など）を検知、捕捉、変換し、機械読み取り可能な生体データ（上記 7.2.1.d) 項を参照）に変換する。これらのデータは、認識目的に使用される個人の顕著な特徴の数学的表現である画像またはテ ンプレートという異なる形態で利用可能である。バイオメトリクス認識技術は、検証および識別の目的で使用される。[185]

(298) AI 法第 3 条 41 項によれば、RBI システムは以下の通りである。

> *[*個人の生体データと参照データベースに含まれる生体データとの比較を通じて、通常、離れた場所にいる自然人を、本人の積極的な関与なしに識別することを目的とした *AI* システム。

(299) この定義は、バイオメトリクス認識システムの識別機能のみを対象としており、関係者 の積極的な関与がない（すなわち、積極的な参加がない）ことを意味し、その結果、通常、 遠距離にいる人の特徴が捕捉される。識別性能のために、捕捉された生体データは、参照データベース（例えば、容疑者の顔画 像またはテンプレートを含む犯罪者データベースなどのリポジトリ）に既に格納されている生体 データと比較される。

a) 本人確認のみを目的とする

(300) バイオメトリクス識別」の概念は、AI 法第 3 条 35 項で次のように定義されている。

> 自然人の識別を確立する目的で、その個人の生体データをデータベースに保存されている個人の生体データと比較することによって、身体的、生理的、行動的または心理的な人間の特徴を自動的に認識すること。

(301) 前文 15 AI 法はさらに、そのような人間の特徴は以下のようなものであることを明確にしている。

[185]ISO/IEC 標準 2382-37:2022 情報技術-用語集、バイオメトリクス認識、用語 37.01.03 でバイオメトリクスコミュニティによって定義されている。

顔、目の動き、体型、声、韻律、歩行、姿勢、心拍数、血圧、匂い、打鍵の特徴などである、

(302) 自然人を追跡するために使用される AI システムも、例えば、容疑者がどの方向に逃げるかを確認するために、バイオメトリクス識別の定義に含めることができる。これは、AI 法第 5 条(1)（h）（iii）が犯罪容疑者の居場所を特定することを認めていることから結論づけられる。位置特定は、人が尾行されているときに可能である。

(303) バイオメトリクス検証に使用されることを意図した AI システムは、AI 法第 5 条(1)(h)の禁止事項の範囲外となる。[186]生体データ検証（または認証）は、センサーに提示されたデータと、スマートフォン、パ スポート、ID カードなどのデバイスに保存されている、以前に記録された別のデータ セットを比較することからなる。バイオメトリクス検証の目的は、特定の人物が本人であることを確認することである。

バイオメトリクス検証の例としては、電子ゲートでスキャンされた旅行者の顔とパスポートに含まれる顔画像の比較がある。

***b)* 遠隔性**

(304) AI 法第 3 条(41)によれば、遠隔性とは、生体データが参照データベースに含まれる生体データと比較されることによって、通常、遠隔地において、本人の積極的な関与なしに個人を識別 する生体システムの能力を意味する。

(305) サービスへのアクセス、機器のロック解除、または敷地へのセキュリティ・アクセスのみを 目的とする、自然人の身元を確認するためのバイオメトリクス・システムの使用は、「遠隔」の 概念から除外される（AI 法の前文 15）。この方式は、たとえばアクセス管理に使用される。[187]

例えば、顔スキャン技術を使って制限区域（例えば発電所構内）に入るために顔識別システムが展開される。このシステムは、入口カメラに提示された個人の顔を、建物への立ち入りを許可された人の参照データベースに含まれる参照画像と比較する。

(306) AI 法の前文 17 は、この禁止範囲からの除外は、多数の個人の生体データの処理に使用され る可能性がある RBI システムと比較して、そのようなシステムが自然人の基本的権利に及ぼす影 響は軽微である可能性が高いという事実によって正当化されることを明確にしている。 前文ではさらに、RBI システムは、一般的に、自然人の積極的な関与なしに自然人の識別を 大幅に容易にするために、複数の人物またはその行動を同時に認識するために使用されることを 明確にしている。積極的な関与のためには、本人が カメラの存在を知らされるだけでは不十分であり、積極的な参 加を促す方法で設置されたカメラの前に積極的かつ意識的に足を踏み入れる必要がある。

例えば、

[186]前文 17 AI 法。

[187]例えば、Ross A, Jain AK (2015) 'Biometrics, Overview' in Li S.Z. and Jain A.K. (eds) Encyclopedia of Biometrics, (1st ed., Springer Science, New York), pp.289-294.Springer Science, New York）, pp.289-294.

- 地下鉄駅の壁や天井に設置された監視用カメラに使用される RBI システム。このようなシステムは、遠隔という条件を満たしている。

- バイオメトリック・メトロ・チケットなど、地下鉄駅へのアクセスに使用されるシステムは、人が積極的に関与し、アクセスを得るために意識的にバイオメトリック・センサーに近づくものであるが、この条件を満たしていない。

(307) 非接触）指紋、歩行、音声、DNA、キーストローク、その他の（バイオメトリクス）行動信号を処理するバイオメトリクス認識システムも、RBI システムを構成する可能性がある。[188]

例えば、

- 音声バイオメトリック技術システムは、話している人を識別するために展開することができる。その後、マイクロホンが生体サンプルを収集する。

- CCTV を介して歩行認識システムを使用することもでき、ビデオは自動的に、以前に撮影されたテンプレートとの一致がチェックされる。

- キーストロークの生体認証技術は、不正なメッセージを入力した人物を特定するために使用される可能性がある。

これらのシステムが RBI システムの例として挙げられているからといって、AI 法第 5 条で禁止されているわけではない。

(308) 個々の警察官が使用できる団体カムの場合、例えば数百人の参加者がいるデモの最中にターゲットを絞らずに撮影することは、遠隔性の条件を満たしているとみなされる。

c) 参照データベース

(309)識別は、比較目的の生体データを含む参照データベースなしには不可能である。したがって、参照データベースの存在は、識別目的の比較を実行するために**不可欠である**。[189]

たとえば行方不明者の場合、シェンゲン情報システム（[190] ）のデータベースを、（運用開始後は）顔認識目的の参照データベースとして使用することができる。

9.2.2. リアルタイム

(310) リアルタイムとは、システムが生体データを取得し、さらに処理することを意味する。

[188]EDPB-EDPS, Joint Opinion 5/2021, p. 11; Council of the European Union, 'Opinion of Legal Service', 12302/22, 12 September 12 2022, paragraph 33, and Recital 15 AI Act.

[189]前文 34 AI 法。

[190]行方不明者に関するアラート（SIS II 決定の第 32 条）；第 2 世代シェンゲン情報システム（SIS II）の確立、運用、使用に関する 2007 年 6 月 12 日の理事会決定 2007/533/JHA

即座に、ほぼ即座に、またはいかなる場合にも大幅な遅延なしに」。[191]すべての処理段階、すなわち生体データの取得、比較、および識別は、同時またはほぼ同時 に行われるが、RBI システムの遡及的使用によって禁止が回避されるのを避けるために、「限 定された短い遅延」を含めることができる。[192] 大幅な遅延なく」という概念は AI 法では定義されていないため、ケースバイケースでアセスメントする必要がある。 リアルタイムの本人確認や事後の本人確認に使用されるデバイスは、異なる機能を持つ同一 のものとなりつつあるため、その区別は時間的なものとなる。一般的に言って、遅延は、少なくとも生体データが採取された場所から個人 が立ち去った可能性が高い場合に顕著となる。

(311) リアルタイム・システムは一般に、ある場所で迅速な反応を促すために使用されるものであり、人物を遡及的に識別するために使用されるものではない。監視対象者の動きを追跡し、監視する手段をシステムのユーザーに提供する。

a) AI システムがコンサート会場の入場者全員を選別する：リアルタイム RBI

b) コンサートに入場するすべての客を撮影するシステムがある。コンサート会場でインシデントが発生する。コンサート終了後、犯人を特定するため、識別システムが映像資料を操作する：ポスト RBI である。

(312) 法執行当局が携帯端末を通じて密かに人物の写真を撮影し、即座に検索するためにデータベースに提出する場合、状況によっては AI 法第 5 条(1)（h）の禁止事項に該当する可能性がある。

9.2.3. 公共のアクセス可能なスペースでは.

(313) AI 法第 3 条(44) AI 法は、**公にアクセス可能な**空間とは、不特定多数の自然人がアクセス可能な、公有または私有の物理的空間であると定義している。

(314) 前文 19 AI 法は、このようなスペースを特徴づけるいくつかの要素を挙げている：

- チケットの購入や交通機関の名義、事前の登録、一定の年齢など、潜在的な定員やセキュリティの制限とは無関係に、不特定多数の人がアクセスできること。無施錠のドアからスペースにアクセスできる可能性があっても、（アクセスを制限する標識のような）反対を示唆する表示や状況があれば、そのスペースが公にアクセス可能であることを意味しない。さらに、ある空間への立ち入りを、公共の安全やセキュリティに関連する、法律で定められた特定の人（ ）に制限したり、その空間に対する関連権限を持つ人の決定によって制限したりすることもできる。

[191]前文 17 AI 法。

[192]AI 法第 3 条 42 項

例えば、パブリック・アクセス可能なスペースは原則的に存在する：

- 参加者が入場料を支払うコンサート会場。
- 50 歳以上の参加者を対象とした見本市が開催されるイベント会場。

たとえ門の施錠が解除されていたとしても、数軒の住宅からなるフェンスで囲まれた住宅地の門のような、門によって閉鎖された空間は、通常、公にアクセス可能な空間とはみなされない。対照的に、開門時間があり、開門時間中の立ち入りが制限されていないゲートのある住宅内の公園は、一般に、開門時間中は公にアクセス可能な空間となり、開門時間外は閉ざされた空間となる。

- 所有権の非関連性。つまり、パブリック・アクセス可能な空間とみなされるためには、その空間が公共の所有物である必要はない。

例えば、民間事業体が所有するスペース、公共事業体が所有するスペース、公共事業体が所有し民間事業者が管理するスペースなど、スペースの性質に影響を与えることなく利用することができる。

- そのスペースが使用される特定の活動はない。公にアクセス可能なエリアは、必ずしも公共サービスに関連したスペースとは限らない。さらに、公共サービスに関連するスペースには、一般に公開されていないスペース、すなわち自治体に勤務する公務員のオフィスが含まれる場合もある。

例えば、公にアクセス可能な空間は、商店、レストラン、カフェな どの商業用、銀行、専門家活動（会計士事務所だけでなく医師事務所も含む）、ホスピタリティ（ホテ ルなど）などのサービス用、プール、ジム、スタジアムなどのスポーツ用、バス、地下鉄、鉄道の駅、空港、交通手段などの運輸用、映画館、劇場、美術館、コンサ ートホール、会議場などの娯楽用、あるいはレジャー用に使用することができる。バス、地下鉄、鉄道駅、空港、交通機関などの交通機関、映画館、劇場、美術館、コンサートホール、会議場などの娯楽施設、公道や広場、公園、森林、運動場などのレジャー施設。[193]

(315) 以下のスペースは、AI 法第 5 条(1)第 1 号（h）の意味における公共のアクセス可能なスペースには該当しない：

- オンラインスペースは、AI 法第 3 条 44 項にいう物理的空間を構成しないためである。

例えば、チャットルーム、ソーシャルメディア、オンラインプラットフォームなどは、そのため禁止事項の範囲から除外される。

[193]前文 19 AI 法。

- 工場、会社、職場など、関係する従業員やサービス提供者のみがアクセスすることを意図しているため、関係する従業員やプロバイダのみにアクセス管理または制限されている、限られた人数がアクセスすることを意図した特定のスペース 。[194]

例えば、バッジでアクセスできる職場は原則的にパブリック・アクセス可能な空間とはみなされないが、アクセス管理のないオフィスはパブリック・アクセス可能な空間とみなされる可能性がある。

- **刑務所や国境**管理は、一般人がアクセスできる空間ではない。[195]

(316) 例えば、国境を越える地点は公共のアクセス可能な空間ではないが、国境を越える地点に通じる道路やその周辺の森林は通常、公共のアクセス可能な空間である。

(317) 一部のスペースは二重の機能を持ちうる。例えば、空港は全般的にその共用部分に関しては公的にアクセス可能な空間とみなされるが、国境管理専用のエリア（税関職員が立ち、パスポートや ID チェックが行われる場所）は禁止の範囲から除外される。

(318) AI 法 19 条前文で明確にされているように、あるスペースが公衆にとって利用しやすいかどうかの分析は、ケースバイケースの分析に基づいて行われるべきである。

9.2.4. 法執行目的の場合

(319) AI 法第 5 条(1)(h)の禁止は、法執行活動を行う事業体、認可、団体に関係なく、法執行目的のための RBI システムの使用に適用される。

(320) 法執行とは、AI 法第 3 条 46 項において、「犯罪の予防、捜査、探知もしくは訴追、または刑事罰の執行のために、法執行当局によって、または法執行当局に代わって行われる活動であり、公共の安全に対する脅威の保護および防止を含む」と定義されている。これらの目的は、LED 第 1 条に記載されている目的と同じである。[196]したがって、LED に関連するこれらの目的の解釈は、AI 法で使用されている「 () 法の執行」の概念を解釈する目的にも関連する可能性がある。

(321) 法執行目的は、犯罪の捜査、検知、起訴からなる。また、犯罪が実際に行われる前に、 公共の安全に対する脅威の保護と防止を含む、犯罪の防止に関する活動も含まれる。例えば、警察は犯罪予防の観点から「デモ、大規模なスポーツイベント、暴動における強制措置」をとることがある。[197]最後に、これらの活動は、刑の執行など刑罰の執行からなる。

[194]前文 19 AI 法。

[195]前文 19 AI 法。別の文脈では、国境管理は、規則（EU）2016/399（シェンゲン国境コード）に従い、またその目的のために、その国境を越える意図またはその国境を越える行為のみに対応して、国境で実施される活動と定義されている。これは、いわゆる国境地域は含まず、国境の両側で最大 50 キロメートルまで拡大することができる。

[196]法執行当局の外国の活動には、行政業務（人事など）を行う場合など、LED の範囲から除外されるものがあるが、これらの活動は法執行の枠外で行われる。これらは GDPR に該当する。GDPR の前文 19 を参照のこと。

[197]前文 12 LED。

(322) AI 法第 3 条 46 項によれば、法執行活動は、法執行当局または法執行当局に代わって行われる。法執行機関はさらに AI 法 3 条 45 項で、LED における国家認可当局の定義と同様に定義されている。[198]この定義は、法執行当局および委託された団体または事業体（私人である場合もある）を対象としている：

(a) 刑事犯罪の予防、捜査、検知もしくは訴追、または刑事罰の執行（公共の安全に対する脅威の保護および予防を含む）を管轄する公的機関。

たとえば、警察当局や刑事司法当局（検察官など）が法執行を行う場合、そのような公的当局が含まれる。

(b) 刑事犯罪の予防、捜査、摘発、訴追、刑事罰の執行を目的として、加盟国の法律によって公権力および公権力の行使を委託されたその他の団体または事業体（公共の安全に対する脅威の保護および予防を含む）；

(323) AI 法に基づき、他の事業体、団体または個人は、上記の目的のために、公権力および公権力を委任した加盟国の法律により委託された後、法執行活動を行うことができる。

(324) のために」とは、法執行当局が法執行活動（またはその一部）の遂行を、私人を含む他の事業体または個人に委任したこと、あるいは特定の場合に、法執行活動を支援するために行動するよう他の事業体または個人に要請したことを意味する。いずれの場合も、法執行当局は、すべての主要な側面について指示し、他の事業体を監督しなければならない。この要件は、「人のために」行動するという概念に内在するものだからである。

他の団体への業務委託には、例えば以下のようなものがある、 - 公共交通機関は、法執行当局の指示・監督のもと、公共交通網の安全確保を要請される。 - 法執行当局の要請を受けたスポーツ連盟は、その指示・監督のもと、スポーツイベントの警備を行う。 - 法執行当局の指示・監督の下、「特定の事案における特定の犯罪に対抗する」ために、法執行当局から特定の行為を行うよう要請された銀行。 これらの事業体は、法執行当局に「代わって」行動するため、「法執行のため」の定義に該当する。これらの事業体が犯罪（詐欺やマネーロンダリングなど）を検知し、これに対抗する際に「自らのために」行動するのであれば、AI 法第 5 条(1)（h）の禁止事項には該当しないと考えられる。

[198]第 3 条（7）LED。

(325) その他の団体や事業体が特定の法執行任務を委託された場合にのみ、その活動は「法執行」の定義に該当する。

9.3. 禁止事項の例外

(326) AI 法は、公共のアクセス可能な空間におけるリアルタイム RBI の法執行目的での使用に対する一般的な禁止に対して、3 つの例外をプロバイダとして規定している。第 5 条(1)(h)(i)から第 5 条(1)(i)までである。

(iii) AI 法は、リアルタイム RBI が認可される 3 つの目的を網羅的に列挙しており、AI 法第 5 条(2)～(7)は、そのような認可の条件と保護措置を定めている。AI 法第 5 条(1)(h)(i)～(iii)は、それ自体、一般にアクセス可能な空間における RBI システムのリアルタイム利用の法的根拠を構成するものではない。むしろ、特に第 5 条(2)(7)AI Act の要件を満たす加盟国の国内法のみが、第 5 条(2)AI Act が規定するリアルタイムの RBI の使用を認めることができる。従って、1 つ以上の目的のためにリアルタイム RBI の使用を許可する加盟国の法律がない場合、そのような使用は 2025 年 2 月 2 日から禁止される。

9.3.1. 根拠と目的

(327) AI 法第 5 条(1)(h)(i)～(iii)に規定された目的は、法執行目的のために特定の AI や捜査ツールの使用を認めることである。 これらの目的は以下の通りである：

(i) (i)3 つの特定の重大犯罪の被害者および行方不明者の標的を絞った捜索［保護］;

(ii) 生命や身体の安全に対する差し迫った脅威、またはテロ攻撃の真の脅威の防止［予防］。

(iii) 附属書 II［訴追／捜査］に列挙された特定の重大犯罪の容疑者と犯罪者を特定し、特定する。

(328) こうしたシナリオにおいて、連邦立法府は、社会の安全保障上のニーズと、リアルタイム RBI システムがその対象となる個人の基本的権利にもたらすリスクとのバランスをとってきた。AI 法の前文 33 によれば、公共のアクセス可能な空間における法執行目的のリアルタイム RBI システムの利用が認められる目的は、厳格かつ網羅的で狭く定義されなければならず、基本的権利にもたらされる「リスクを上回る」「実質的な公共の利益」を達成するために「厳格な必要性」がある場合にのみ認められる。AI 法第 5 条(1)(i)～(iii)に列挙されていない、法の執行を目的とした、公共のアクセス可能な空間におけるリアルタイム RBI システムのその他の使用は禁止される。

例えば、万引き犯を特定し、その顔画像を犯罪データベースと比較するために警察がリアルタイム RBI システムを使用することは、AI 法第 5 条(1)（h） （i）～（iii）に列挙された目的のいずれにも該当しないため、禁止されている。

9.3.2. 重大犯罪の被害者と行方不明者を対象に捜索を行う。

(329) AI 法第 5 条(1)(h)(i)によれば、拉致、人身売買、性的搾取の被害者、および行方不明者の捜索のため に、厳密な必要性および AI 法第 5 条第 2 項～第 7 項の条件に従い、法執行目的で公衆がアクセス可 能な空間におけるリアルタイム RBI を使用することが認められている。

a) 3 種類の犯罪の被害者を対象として捜索する。

(330) AI 法第 5 条(1)(h)(i)のシナリオは、法執行当局が 3 つの重大犯罪の被害者を捜索するのを支援しようとするものである。

(331) 標的を絞った捜査には、被害者の位置特定と身元確認が含まれる。*3* 種類の犯罪

(332) AI 法第 5 条(1)(h)(i)に掲げるシナリオの対象となるのは、人間の拉致、人身売買、性的搾取という 3 つの重大犯罪の特定の被害者を対象とした捜索である 。[199]

> 例えば、子供が誘拐され、誘拐犯が子供をある場所から別の場所に車で連れて行こうとしているという具体的な兆候がある場合、警察はその子供を対象とした捜索のためにリアルタイム RBI システムを使用することができるが、子供を特定するための展開範囲と使用期間を定めなければならない。

b) 行方不明者の捜索

(333) 最初のシナリオは行方不明者の捜索もカバーしている。[200]

(334) 行方不明の成人の自発的な失踪が必ずしも捜索の引き金になるとは限らないため、行方不明の子どもと行方不明の成人は区別することができる。行方不明の子どもに関して適用される規則は、加盟国によってかなり異なる。[201]いずれにせよ、AI 法第 5 条(1)(h)(i)は、法執行の目的で行方不明者を捜索するためにリアルタイム RBI システムを使用することのみを認めている。

(335) 成人には失踪する権利があるため、成人の失踪が必ずしも警察の捜索につながるとは限らない。捜索は、その人の法的地位（「被保佐人」）、健康状態（精神疾患）、自殺のメモの有無だけでなく、所持品を持たずに出かけたこととも関連する可能性がある。失踪の状況が懸念されるものであれば、失踪届を警察に提出し、捜索を開始することができる。

(336) 加盟国によっては、行方不明者の捜索は行政手続きの下で行われ、法執行目的ではない。例えば、脆弱性が行方不明であるが、犯罪の疑いやその他の法執行目的がない場合、その人物を捜索す

---

[199]誘拐、人身売買、性的搾取の 3 つの犯罪は、欧州逮捕状（EAW）を発動し、犯罪容疑者または刑に服した者を逮捕し、EAW を発行した国に移送することができる。この 3 つの犯罪は、ほとんどが女性と子どもに関するものであるが、それだけに限定されるものではない。欧州委員会の移民・内務総局によれば、被害者のほぼ 40%が EU 市民であり、そのほとんどが性的搾取目的で人身売買された女性と子どもである。男性の被害者数は 10 年間でほぼ倍増している。彼らは強制労働や強制物乞いのために人身売買されているが、女性と子どものほとんどは性的搾取のために人身売買されている。

[200]行方不明者」は EU レベルでは定義されていない。しかし、「行方不明者分野における国境を越えた警察協力の強化」に関する 2021 年 12 月の理事会結論において、理事会は欧州評議会の勧告 CM/Rec（2009）12 における行方不明者の定義と国内規則の両方を参考にしている。理事会結論（2021）14808/21、パラ 11、4 ページ

[201]欧州委員会、欧州移民ネットワーク、「EU 加盟国は同伴者のいない未成年者の行方不明事例をどのように扱っているか？EMN Inform, 2020

るためにリアルタイム RBI システムを使用することは、法執行目的とはみなされず、したがって GDPR の当該使用に関する規則に該当する。

9.3.3. 生命に対する差し迫った脅威やテロ攻撃の防止

(337) AI 法第 5 条(1)項(h)(ii)は、厳密な必要性および AI 法第 5 条(2)～(7)項に定める条件（自然人の生命ま たは身体の安全に対する具体的、実質的かつ差し迫った脅威、または真正かつ現在もしくは真正 かつ予見可能なテロ攻撃の脅威の防止）を条件として、法執行目的で公共のアクセス可能な空間におけるリアルタイム RBI の使用が許可される第 2 のシナリオを挙げている。

(338)

***a)* 自然人の生命または身体の安全に対する具体的、実質的かつ差し迫った脅威**

(339) 生命への権利を保証する憲章第 2 条の適用により、EU および加盟国は個人の生命を保護しなければならない。AI 法第 5 条(1)(h)(ii)にある、公共のアクセス可能な空間におけるリアルタイム RBI システムの使用を認めるための、生命への脅威に関する規準は、(1) 、(2)実質的かつ(3)差し迫った、(4)自然人の生命または身体の安全への脅威の存在を必要とする。脅威は、自然人一般に関係するものであるため、識別された個人または集団に限定する必要はない。

(340) 前文 33 AI 法は、自然人の生命または身体の安全に対する差し迫った脅威には、重要インフラ（[202] ）に対する差し迫った脅威も含まれることを明確にしている。

> 例えば、[203]
>
> 重要インフラ（発電所、水道、病院など）の深刻な混乱や破壊は、住民への基本的な供給が停止するという深刻な被害（電気や飲料水が長期間にわたって奪われる、特に温暖な気候や寒冷な気候など）がある場合、人の生命や身体の安全に対する差し迫った脅威をもたらす可能性がある。

(341) 何が自然人の生命または身体の安全に対する差し迫った脅威を構成するかは、最終的には加盟国のレベルで、EU 法に従い、特に AI 法第 5 条の重要な要素と根拠を考慮して、国内法に基づいて定義され、評価される。このことは、加盟国が、公共のアクセス可能な空間における法執行目的のリアルタイム RBI の使用禁止の例外を利用するために採択しなければならない法律に規定／言及されなければならない。

(342) 生命または身体の安全に対する**差し迫った**脅威とは、いつ起こるかわからない脅威であり、「**直ちに行動を起こす**」必要がある□[204]□身体の安全に対する**実質的な**脅威とは、深刻な身体的傷害に関するものである。

[202]指令 2022/2557 の第 2 条 4 項に定義されている。

[203]前文 33 AI 法。

[204]規則 2023/1543 の前文 37。

(343) 具体的な脅威とは、その脅威が明確に定義され、個別化され、具体的であることを意味する。

> 例えば、ある元学生が、かつての同級生数人に復讐するために、かつて在籍していた大学で致命的なテロを計画しているという情報が警察に入ったとする。警察は、襲撃の切迫性、標的となる学校、計画実行のために彼が使用する予定の武器に関する情報を入手する。

(344) 具体的な脅威は意図的である必要はない。意図的でない行動も、生命や身体の安全に対する脅威となりうる。

***b* 真正かつ現在、または真正かつ予見可能なテロ攻撃の脅威。**

(345) AI 法第 5 条(1)（h）(ii)に記述されている第 2 のシナリオのこの部分は、いくつかの要素で構成されている。すなわち、テロ攻撃の脅威の存在と、その脅威の特徴であり、真正かつ現存、または真正かつ予見可能でなければならない。

テロ攻撃の脅威

(346) 脅威の存在と深刻さに関するアセスメントは、国家安全保障を守るために講じられる措置の実際の状況を評価する際、より具体的にはテロ攻撃の場合に、国家レベルで行われる。**テロの脅威レベルは国レベルで定義さ**れ、加盟国によって異なる。[205][206][207]例えば、オランダは 5 段階、ベルギーは 4 段階、フランスは 3 段階、スウェーデンは 5 段階である。[208]しかし、第 5 条(1)(h)(ii)で使用されている「真正かつ現在の、または真正かつ予見可能な脅威」という概念は、連邦法の独立した概念であるため、原則として各国の定義とは無関係にアセスメントされるべきである。脅威はテロリズム全般に関するものではなく、特にテロ攻撃の脅威に関するものである。

脅威の特徴：真正かつ現在、または真正かつ予見可能である。

(347) 法執行の目的で、一般にアクセス可能な空間におけるリアルタイム RBI システムの使用を許可するために脅威が到達すべき深刻さの閾値は、国家安全保障、特にテロ攻撃に対する安全保障を目的としたデータ保持と乗客名記録措置に関する CJEU の判例法に触発されたものである。CJEU によれば、このような文脈では、「国家安全保障に対する脅威は、真正かつ存在するか、少なくとも予見可能でなければならない。[209]

予防

---

[205] https://www.government.nl/topics/counterterrorism-and-national-security/risk-of-an-attack-threat-level

[206] https://cuta.belgium.be

https://crisiscenter.be/en/risks-belgium/security-risks/terrorism-and-extremism

[207] https://www.sgdsn.gouv.fr/vigipirate#

https://www.sgdsn.gouv.fr/files/files/Vigipirate/20160130-np-sgdsn-pse-tackling-terrorism-together.pdf [208]

[208] https://www.krisinformation.se/en/hazards-and-risks/terrorism

[209] 2022 年 9 月 20 日、スペースネット、C-793/19（合議例 C-793/19、C-794/19）、ECLI:EU:C:2022:702、パラグラフ 93 の司法裁判所の判決。

(348) 第 5 条(1)(h)(i)および第 5 条(1)(h)(iii)の AI 法に反して、第 5 条(1)(h)(ii)に記載されたシナリオでは、リアルタイ ム RBI の使用が具体的な人物を特定または識別するために許可されるとは明記されていない。その目的は、特定の脅威の防止である。したがって、このシナリオは、特定の脅威を検知し追跡するためのリアルタイム RBI の使用も対象とすることができる。移動中のテロリスト」、すなわち、同じ脅威に関連する複数の人物が、テロ攻撃を計画しているという具体的な兆候があるが、その場所が明確でない場合である。

公園でのテロ攻撃を防ぐリアルタイム RBI ある人物が、通常テロ攻撃やテロ集団に関連する暴力的な過激派スローガンを叫びながら、ナイフで人を襲うために公園を走り回っていると警察に通報される。 加盟国が AI 法第 5 条(1)(h)(ii)のシナリオでリアルタイム RBI の使用を認可している場合、法執行当局は、AI 法第 5 条(2)～(7)の他の条件が満たされていれば、攻撃を防ぐために、公園およびその周辺にいる人物を特定し、居場所を突き止めるためにリアルタイム RBI を使用することができる。

9.3.4. 特定の犯罪の容疑者の位置特定と識別

(349) AI 法第 5 条(1)(h)(iii)は、「附属書Ⅱに言及される犯罪であって、当該加盟国において最高 4 年の拘禁刑又は拘禁命令により処罰されるものについて、犯罪捜査若しくは訴追を行い、又は刑事罰を執行する目的で、犯罪を犯したと疑われる者の位置特定及び特定」のために、公共のアクセス可能な空間における RBI のリアルタイムの使用を認めている。

AI 法附属書Ⅱは、前述の目的のためにリアルタイム RBI の使用が許可される重大犯罪のリストを網羅している。これらの犯罪は以下の通りである： - テロリズムだ、 - 人身売買だ、 - 児童の性的搾取、児童ポルノ、 - 麻薬や向精神薬の不正取引、 - 武器、軍需品、爆発物の不正取引、 - 殺人、重傷傷害、 - 人間の臓器や組織の不正取引、 - 核物質や放射性物質の不正取引、 - 誘拐、違法な拘束、人質取り、 - 国際刑事裁判所の管轄下にある犯罪、

- 航空機や船舶の不法な占領、

- レイプ、

- 環境犯罪、

- 組織的または武装強盗、

- 妨害工作、

- 上記の犯罪の 1 つ以上に関与する犯罪組織への参加。

a) 位置特定と識別

(350) 加盟国は、犯罪捜査の実施、犯した犯罪に対するその者の訴追、または既存の刑の執行のために、刑事犯罪の容疑者の所在を突き止め、識別することを目的とする法執行のために、公にアクセス可能な空間におけるリアルタイム RBI の使用を許可することができる。

b) 容疑者と加害者

(351) AI 法第 5 条(1)(h)(iii)は、被疑者と加害者という 2 つのカテゴリーを対象としている。被疑者とは、刑事犯罪を犯したと信じる重大な根拠があり、その人が犯罪に関与しているという十分な証拠がすでに収集されている人である。加害者とは、刑事犯罪を犯したとして告発され、または有罪判決を受けた者をいう。附属書Ⅱに記載された犯罪の共犯者の所在を突き止めたり、特定したりする場合にも、同じ条件（附属書Ⅱに記載された犯罪、最高刑が 4 年以上）が適用される。

c) 重大犯罪リスト

(352) 法執行の目的で、公共のアクセス可能な場所でリアルタイム RBI システムを使用することが正当化されるのは、重大犯罪のみである。

(353) 附属書Ⅱの AI 法に記載されている最初の 5 つの犯罪は、TFEU 第 83 条に記載されている「ユーロ犯罪」と同じである。[210]そのうちのいくつか（誘拐、核物質や放射性物質の不正取引など）は、テロリズムと関連している可能性がある。[211]

(354) 附属書Ⅱに列挙された犯罪はすべて、容疑者または加害者に対する欧州逮捕状（「EAW」）の発付の引き金となる可能性があるが、これらの重大な犯罪の一つについて容疑者の所在を突き止め、特定するためにリアルタイム RBI を使用する場合、EAW が発付されている必要はない。

(355) さらに、この目的のために実時間 RBI を使用するには、それぞれの犯罪が、当該加盟国において、少なくとも 4 年以上の拘禁刑または拘禁命令によって罰せられるものでなければならない。

[210]欧州刑事警察機構の優先事項だ。

[211]前文 33 および指令 2017/541 第 3 条のテロ犯罪の定義を参照のこと。

ある都市の賑やかな祭りの期間中、警察当局はライブ顔認識技術を展開して祭り周辺を監視し、違法薬物密売や性犯罪で逮捕状が出ている指名手配者を特定する。フェスティバルのさまざまな入口で、警察は携帯カメラの前を通過する人々のライブ・ビデオ映像を使い、指名手配者の顔の監視リストと比較する。

まず、犯罪の種類についてだが、RBI は違法薬物売買の場合に使用できる。しかし、性的犯罪は、児童の性的搾取、児童性的虐待の材料、強姦に関するものでない限り、犯罪リストには含まれていない。警察は、リアルタイム顔認識技術を、指名手配中の犯罪者を発見し、路上から排除することを目的として、対象を絞らずに大々的に展開することは認められていない。

警察が、麻薬密売で欧州逮捕状が出されている指名手配中の人物の写真付き人相を入手し、その人物がフェスティバルの会場に現れると信じるに足る理由がある場合は異なる。このような状況では、対象となる個人を特定するためにリアルタイムの顔認識技術を展開することは、AI 法第 5 条(1)(h)(iii)の対象となる可能性がある。

クリスマスマーケットで 12 人の死者を出すテロ事件が発生した後、警察は犯人を特定し、逃走先を確認するためにリアルタイムの顔認識技術を使用した。その際、近くの駅や、テロ直後にそこから出発する列車の行き先駅のリアルタイム顔認識技術も使用している。テロ攻撃の場合、このような利用は AI 法 5 条 1 項 3 号で認められる。

(356) 第 5 条(1)項(h)(i)と第 5 条(1)項(h)(iii)AI 法の間の関連は、第 5 条(1)項(h)(i)AI 法に記載されたシナリオの対象となる犯罪につい ても考えられる。リアルタイム RBI システムは、被害者や行方不明者を発見するために展開されるかもしれないが、これらのシステムは、人身売買、（附属書Ⅱに記載されている）児童に関する限りにおける性的搾取、誘拐（附属書Ⅱ AI 法第 5 条(1)(h)(i)に記載されている誘拐が、附属書Ⅱ AI 法に記載されている誘拐に該当する限りにおいて）の加害者や容疑者を発見し、特定するためにも使用することができる。リアルタイム RBI システムは、第 5 条(1)（h）（ii）の範囲に入る脅威を防止するために使用することができ、その脅威が現実化した場合、それらのシステムは、移動中の[「] 犯人を特定／所在を特定するために使用することができる。

10. AI 法第 5 条(2) (7) - 例外に関する保護措置と条件

10.1 対象となる個人とセーフガード（AI 法 5 条 2 項

***AI* 法第 *5* 条 *2* 項はプロバイダを規定している：**

第 1 項第 1 号の(h)に言及されるいずれかの目的のための法執行の目的で、公共のアクセス可能な空間における「リアルタイ ム」遠隔バイオメトリクス識別システムの使用は、具体的に対象となる個人の身元を確認するためにのみ、同 号に規定される目的のために展開されるものとし、以下の要素を考慮しなければならない：

(a) 使用される可能性のある状況の性質、特にシステムが使用されなかった場合に生じるであろう損害の重大性、蓋然性、規模；

(b) システムの使用がすべての関係者の権利と自由に及ぼす影響、特にその影響の重大性、蓋然性、規模。

さらに、本条第 1 項第 1 号(h)で言及されるいずれかの目的のための法執行の目的で、公共のアクセス可能な空間における「リアルタイム」遠隔バイオメトリクス識別システムの使用は、その使用を許可する国内法に従い、特に時間的、地理的、個人的制限に関して、その使用に関して必要かつ相応の保護措置および条件を遵守しなければならない。公にアクセス可能な空間における「リアルタイム」遠隔バイオメトリクス識別システムの使用は、法執行機関が第 27 条に規定する基本的権利影響アセスメントを完了し、第 49 条に従って EU データベースに登録した場合に限り認可されるものとする。ただし、正当に正当化された緊急の場合には、不当に遅延することなく登録が完了することを条件に、EU データベースに登録することなく、そのようなシステムの使用を開始することができる

(357) AI 法第 5 条(1)(h)(i)から(iii)に列挙された目的の 1 つに対するリアルタイム RBI システムの使用は、AI 法第 5 条(2)から第 5 条(7)に詳述されている一定の保護措置と条件の対象となる。

(358) 第一に、法執行の目的で、公共のアクセス可能な空間におけるリアルタイム RBI システムの使用は、「特に標的とされた個人の身元を確認する」場合にのみ許可される。この最初の条件は、事態の深刻さと、システムを使用しないことによる弊害と、個人の権利と自由に対する技術の影響とのバランスをとることを目的としている。これは、リアルタイムの RBI を展開するために個人をターゲットにすることで、集団監視を避けることを目的としている。その結果、法執行を目的とした、公共のアクセス可能な空間におけるリアルタイム RBI システムの展開は、対象となる個人に対してのみ認可されるべきである。

(359) 識別」ではなく「身元確認」という表現の使用は、 無差別監視のリスクを制限する基本的権利の追加的保護として意図されており、AI 法第 5 条(1)(h)の意味における個人の識別は対象を絞らなければならないことを意味する。この表現は、法執行当局が、AI 法第 5 条(1)（h） （i）に列挙された犯罪の被害者である、または AI 法第 5 条(1)（h） （ii）もしくは同条 1 項（h） （iii）に列挙されたシナリオのいずれかに関与していると信じるに足る理由があるか、またはそうであると知らされた特定の個人を捜索するためにのみ、リアルタイム RBI の使用を開始できることを意味すると理解すべきである。これは実際には、リアルタイムで収集されたデータと参照データベースに含まれるデータの比較を意味する。AI 法第 5 条(1)(h)(ii)に記載されたシナリオにおけるリアルタイム RBI システムの使用、および AI 法第 5 条(1)(h)(iii)の意味における犯罪捜査の実施に関しては、法執行 当局は、システムを使用する前に、検索している個人の身元を必ずしも知る必要はない。特定の時間および場所におけるテロリスト・グループによるテロ攻撃計画（誰がその計画を実行するかは知らないが）に関する事実の兆候および情報を持っている場合、法執行当局がテロリスト・グループの一員である個人の生体データを含む参照データベースを構成していれば、RBI システムを使用してテロリスト・グループから犯罪者を特定することができる。AI 法第 5 条(1)(h)(i)～(iii)に記載されている 3 つのシナリオすべてにおいて、「身元を確認する」 ことは、問題の人物の地域特定も含むことができる。

(360) 第二に、システムを使用する前に、その使用の可能性を生じさせる状況の性質、特に、システムが使用されなかった場合に生じる自然人、社会及び法執行目的に対する危害の重大性、蓋然性及び規模を、システムの使用が関係者の権利及び自由に及ぼす結果、特に、それらの結果の重大性、蓋然性及び規模に照らして評価すべきである。これには、法執行当局または法執行当局に代わって行動する事業体にとって、より侵入的でない代替的解決策が利用可能かどうかの評価も含まれるべきである。

例えば、警察当局は、一般的なセキュリティ、犯罪防止、過密状態の懸念に基づき、街頭でリアルタイムの顔認識システムを使用することを禁じられている。なぜなら、それはすべての人の常時監視と監視を伴うものであり、時間に制限がなく、したがって、AI 法第 5 条(1)（h）に規定された禁止の例外の規準を満たさないからである。

(361) **深刻さ**」の基準は、起こりうる損害と結果に関連して適用されるもので、基本的権利への干渉の程度にばらつきがあることを意味し、これは比例原則と関連している。[212]基本的権利への干渉に関しては、ある干渉は他の干渉よりも深刻であるとみなされる。

(362) **規模**」の基準とは、特に、干渉の影響を受ける人の数とカテゴリーを指す（子どもや脆弱な立場にある人、社会から疎外された人を含む）。

(363) 最後に、「**確率**」とは、ある事象が発生する可能性のことである。

(364) 危害と結果の重大性、規模、蓋然性の認可はすべて、法執行機関に義務付けられている基本的権利影響アセスメントの一部であるべきだ（下記参照）**。**そのアセスメントはケースバイケースで結論が出される。

(365) 第三に、RBI のリアルタイム利用は、地理的範囲、利用時間、対象者を明確に限定すべきである。これは、RBI システムが厳密に必要な場合にのみ使用されるようにするためである。

(366) **地理的制限については**、「客観的かつ非差別的な要因」に基づき、1 つまたは複数の地理的地域を対象とすることができる。バイオメトリクス識別の場合、このことは、地理的制限が、事象が発生することが示唆さ れる、明確に画定された境界線に適用されることを意味する。このような境界線は、通常であれば、都市全体や国全体であってはならない

(367) もう一つの保護措置は、措置の**個人的範囲**、すなわち**関係者のカテゴリーを**定めることに関するもの**である。**これによって、インシデントを示すことなく、無差別に個人を特定することは排除される。

(368) 最後に、**期限は**厳密に必要なものに限られるが、適用される規則に従い、必要な場合には延長することができる。したがって、リアルタイム RBI システムの使用は、無期限または曖昧な期間とすることはできない。その期間は、RBI システムの使用につながる具体的な兆候に照らして決定される必要がある。

[212]2018 年 10 月 2 日の司法裁判所の判決、*Ministerio Fiscal*, C-207/16, ECLI:EU:C:2018:788, paragraph 55 では、「アクセスは、問題となっている基本的権利への干渉の重大性に比例したものでなければならない」としている。

(369) 第 4 に、リアルタイム RBI システムを展開する法執行当局は、展開前に基本的権利影響アセスメント（FRIA）を実施し、EU のデータベースにシステムを登録しなければならない（正当に正当化された場合を除く）。

10.1.1. 基本的人権の影響アセスメント

(370) AI 法第 5 条第 2 項を適用して実施される FRIA は、AI 法第 27 条に規定された条件を遵守しなければならない。この規定は、リスクの高い AI システムに適用される FRIA に関する要件を定めている。

(371) AI 法第 5 条の禁止事項が適用され（2025 年 2 月 2 日以降）、ハイリスク AI システムに関する規定がまだ適用されていない期間（2026 年 8 月 2 日以前）には、AI 法第 5 条(1)(h)の 1 つ以上の例外の恩恵を受ける条件を満たすリアルタイム RBI システムの展開者は、AI 法第 27 条に定める FRIA の要件を実施すべきである。以下の暫定ガイダンスは、高リスクの AI システムに対する義務が適用され、欧州委員会が FRIA のひな型を採用し、AI 法第 27 条に基づく義務についてさらなるガイダンスを提供するまでの期間、法執行を目的とした一般にアクセス可能な空間におけるリアルタイム RBI の使用に関するものである。

(372) FRIA は新しいタイプの影響アセスメントであり、RBI システムを含む特定のリスクの高い AI システムが基本的権利に与える影響を特定することを目的としている。FRIA は説明責任のためのツールである。FRIA は、データ管理者（すなわちパーソナルデータの処理に責任を負う者）が第 27 条 LED、第 35 条 GDPR、第 39 条 EUDPR に基づいて実施しなければならない既存のデータ保護影響アセスメント（DPIA）に取って代わるものではない。

> 例えば、公共のアクセス可能な空間において、（CCTV、AI 顔認識、身体装着カメラなど）自然人の権利と自由に高いリスクをもたらす可能性が高い新技術によって生体データを処理する場合、DPIA を実施しなければならない。

(373) DPIA がパーソナルデータの処理に起因する個人の権利と自由に対するリスクに重点を置くのに対して、FRIA は AI システムが個人の基本的権利に及ぼす可能性のある影響をより一般的に対象とする。したがって、FRIA の範囲は、対象となる活動やアセスメントされる基本的権利の点でより広範である。個人データが AI システムによって処理される場合（RBI システムの場合）、FRIA は、データ管理者としての展開者が実施する DPIA を補完するものでなければならない。[213] 、DPIA ですでに扱われている側面をカバーせず、重複を避ける。本ガイドラインにおける FRIA の分析は、リアルタイムでの RBI の認可使用に限定されており、 AI 事務局がテンプレートを提供するまでの暫定的な期間において、展開者の予備的なガイダンスと して機能することを目的としている。[214]。

(374) AI 法第 5 条(2)に基づく FRIA の実施義務は、RBI システムの展開者に課されるものであり、事業体や団体、またはそれらに代わって行動するいかなる者にも課されるものではない。他の関係者が展

[213] AI 法第 27 条第 4 項

[214] したがって、この分析はリスクの高い AI システム全般のケースをカバーするものではない。

開者／法執行当局に代わって行動する場合は、FRIA が適切に実施されるよう、すべての関連情報をもって FRIA の準備に貢献しなければならない。

(375) 認可されたリアルタイム RBI システムを展開する前に、FRIA を実施しなければならない。

(376) AI 法第 27 条によれば、FRIA は以下の情報を含むべきである：

- RBI の使用と、その使用に関する展開者のプロセスを、意図された使用目的とともに説明すること：

その説明には以下が含まれる： - 展開者の名前； - リアルタイム RBI システムが使用される法執行目的； - 使用される生体データ（顔画像、音声サンプルなど）のソースを含む、生体識別が比較される参照データベースの説明； - システムの基盤となる技術について、その機能を説明するための記述（プロバイダが提供する利用可能な文書とその名称を参照すること）[215]；- リアルタイム RBI が展開される法的根拠。

- 使用期間と使用頻度

許可された例外の 1 つに対するリアルタイム RBI システムの個々の使用は、AI 法第 5 条 3 項に基づき、展開前に司法当局またはその他の独立機関によって認可されなければならない。対照的に、FRIA の場合、プロバイダは意図される使用期間と予想される頻度を一般的に示さなければならない。

- 制度の影響を受ける人およびグループのカテゴリー

AI 法第 5 条(1)（h）の例外規定の目的上、FRIA は以下を区別すべきである： - 対象となる個人は、犯罪の被害者、加害者、容疑者のいずれかである、 - 生体データが参照データベースに含まれる個人、および - RBI システムが展開される周辺地域に存在する人のカテゴリー。 リアルタイム RBI システムの使用は、対象となる個人の基本的権利に影響を与えるだけではない。比較のために生体データが使用される他の個人、通行人、および捜索区域に偶然居合わせた人々の権

[215]高リスクの AI システムに関する規則が発効すれば、EU のデータベースに登録されたシステムの登録番号と、そこに含まれるシステムの利用可能な情報を参照することでこれを行うことができる。

利も影響を受ける。リアルタイム RBI システムによってカバーされる捜索区域の地理的範囲の記述は、システム によって影響を受ける人の数に影響する。

- 影響を受ける人々への具体的な危害のリスク：

(377) 公共のアクセス可能な空間におけるリアルタイム RBI の法執行目的での使用によって影響を受ける可能性のある基本的権利には、特に以下が含まれる：

- 公共空間における匿名性に対する人々の合理的な期待を含む、私生活と家族生活に対する権利；
- RBI システムは、特定の個人を識別するために生体データやその他の個人データ（氏名、識別番号、民族性などの機微データなど）の処理に依存しているため、データ保護の権利がある；
- RBI システムの使用は、思想、良心、宗教の自由、表現の自由、集会・結社の自由を侵害し、個人の権利と自由の完全な行使を妨げる可能性がある
- 効果的な救済と公正な裁判を受ける権利である；
- システムがバイアス（ジェンダー、民族的、人種的バイアスなど）を埋め込み、容疑者や加害者の誤認につながる場合は、非差別的権利を侵害する；
- 人間としての尊厳の権利が、システムの対象であるかのように感じられるのだ；
- 個人に不利な影響を及ぼすいかなる決定も、リアルタイムの RBI システムの出力のみによって下されることはないからである。
- 被害者、行方不明者、容疑者が未成年者である場合、子どもの権利を保障する；
- 行方不明者が出た場合の高齢者の権利。

特定された影響を受ける人または集団に影響を与える可能性のある具体的な危害のリスクをアセスメント するために、FRIA は、それらの人の基本的権利を特定し、影響を受ける可能性のある人を考慮に入れて、 影響の重大性とその規模を含む基本的権利への影響を評価しなければならない。

FRIA のこの部分には、リアルタイム RBI システムの使用が、その目的及びその使用が意図され ている状況を考慮して、必要かつ適切であるかどうかのアセスメントも含まれるべきである。FRIA は、技術文書と、もし入手可能であれば、その技術がテストされ、バイアスと識別的 を防止するために開発された訓練データに基づいて、システムの性能と精度レベルを記述す べきである。

FRIA はまた、リアルタイム RBI システムの使用が、影響を受ける可能性のあるすべての個人、 特に容疑者または加害者、捜索された被害者、および捜索の対象となる公にアクセス可能な空間に 居合わせたその他の個人の基本的権利に及ぼす影響を特定するべきである。システムがこれらの個人の生体データを処理する限りにおいて、私生活および家族生活に対する権利とデータ保護が影響を受ける

が、データ処理活動に関する限り、DPIA の一部としてアセスメントされる。リアルタイム RBI システムの使用に関連するその他の活動およびその他の基本的権利への影響に ついては、FRIA が DPIA を補完する。展開の状況によっては、人間の尊厳、思想・良心・宗教の自由、集会や表現の自由、効果的な救済と公正な裁判を受ける権利、推定無罪と弁護権、子どもの権利など、これらの個人の他の基本的権利が影響を受ける可能性がある。

FRIA に基づくアセスメントは、AI システムの最初の運用開始前に、抽象的なレベルで行われる べきである。リアルタイムの RBI システムが使用される個々のケースにおける使用の影響を決定する、状況に依 存する具体的な考慮事項は、RBI システムの各使用に関する司法当局または他の独立行政当局に よる認可を要請するための個々の要請の中で、さらに詳しく説明されるべきである（下記 10.23.8.3.項参照）。

- 人的な監督対策

AI 法第 5 条 3 項によると、個人に悪影響を及ぼすような決定は、リアルタイム RBI システムの出力のみに基づいてはならない。結果として、FRIA は、システムの運用中に従う手続きと、意思決定プロセスの中で出力がどのように解釈されるかを記述すべきである。手順は、RBI システムの展開に関する指示を提供し、出力を検証し解釈する人間のエージェントの役割 を明確にし、システムを操作するための訓練を提供すべきである。人間による監督を担当する者は、システムがどのように機能するか、またシステムが不調や誤作動を起こしたときにどのように機能するかを理解するために、十分な「AI リテラシー、訓練、認可」[216] 。

AI 法第 14 条および第 26 条に基づく人間の監視・監督に関するその他の考慮事項も関連性があり、説明する必要がある。

- リスク緩和策

展開者は、（差別的措置の回避を含む）人的監視措置の実施にとどまらず、リスクが顕在化した場合の救済措置について、ガバナンス手続きや苦情処理メカニズム（誤認の場合など）を含めて説明すべきである。

10.1.2. 認可 RBI システムの登録

(378) AI 法第 5 条(2)項もまた、法執行目的で公共のアクセス可能な空間で使用されるリアルタイム RBI システムの展開者に対して、AI 法第 49 条に規定される EU データベースにシステムを登録することを義務付けている。ただし、正当に正当化された緊急事態（差し迫った脅威など）の場合は、法執行当局が不当な遅延（ ）なくシステムを認可すれば、登録前であっても展開を開始することができる。不当な遅延とは、「可能な限り速やかに」を意味すると理解すべきである。

そのシステムを使用する前に登録することができなかった緊急事態の状況。 登録がその基準を満たすかどうかは、ケースバイケースで判断する必要がある。先験的に正確な期限を定めることはできない。遅

[216]216 前文 91 AI 法。

延は意図的な行為によって生じたものであってはならない。AI 法第 49 条 4 項によれば、法執行の目的で使用される RBI システムは、データベースの安全な非公開セクションに登録され、限定された情報と、その情報への限定されたアクセスが提供される。

例えば、使用後 24 時間以内に RBI システムを登録するよう法執行当局に要請することは、実弾射撃を想定した場合など、生命への脅威が差し迫った状況でシステムが展開された場合には、合理的な遅延とみなされるかもしれない。

10.2 事前承認の必要性

(379) AI 法第 5 条 3 項は、リアルタイム RBI システムの個々の利用について**事前の承認を**要求し、**そのようなシステムの出力にのみ**基づく自動意思決定であって、不利な法的効果をもたらすものを禁止して

AI 法第 5 条 3 項はプロバイダを規定している：

第 1 項第 1 号、第(h)号および第 2 項の目的のために、公にアクセス可能な空間における「リアルタイム」遠隔バイオメトリクス識別システムの法執行を目的とした各使用は、司法当局または独立行政当局によって認可され、その決定が使用される加盟国を拘束するものとする。ただし、正当に正当化された緊急事態においては、不当な遅滞なく、遅くとも 24 時間以内に当該認可を請求することを条件として、当該システムの使用を認可なしに開始することができる。当該承認が拒否された場合、当該使用は直ちに中止され、当該使用の結果および出力と同様に、すべてのデータは直ちに破棄され、削除されるものとする。

管轄の司法当局または決定を拘束する独立行政当局は、客観的証拠または提示された 明確な兆候に基づき、当該「リアルタイム」遠隔バイオメトリクス識別システムの使用が、 要請で特定された第 1 項第 1 号の(h)項に規定される目的の 1 つを達成するために必要であり、かつ、そ れに釣り合うものであり、特に、期間、地理的範囲および個人的範囲に関して厳密に必 要なものに限定されることに納得する場合にのみ、認可を与えなければならない。当該認可当局は、請求を決定する際、第 2 項にいう要素を考慮しなければならない。個人に不利な法的効果をもたらす決定は、「リアルタイム」遠隔バイオメトリクス識別シス テムの出力のみに基づいて行うことはできない

10.2.1. 目的

(380) 公共のアクセス可能な空間における「リアルタイム」RBI システムの法執行目的での利用に事前の認可（「事前認可」）を必要とする目的は、そのような目的での利用が想定されるかどうかについてのアセスメントと判断の必要性である：

- 第 5 条(1)（h）（i）～（iii）に列挙された目的のいずれか 1 つ、すなわち、特定の被害者の捜索、特定の脅威の防止、または犯罪者の所在の特定もしくは特定を目的とする目的を達成するために必要かつ相当なものである；

そして

- 期間、地理的および個人的な範囲に関して厳密に必要なものに限定される。

(381) これらの要求の結果、法執行の目的で一般にアクセス可能な空間にリアルタイム RBI システムを展開する前に、二重の必要性と比例性のアセスメントが行われるべきである。第一に、AI 法第 5 条第 2 項が要求するように、FRIA を実行する際に利用者によってアセスメントがなされるべきである。第二に、AI 法第 5 条第 3 項に従い、司法当局または独立行政当局は、憲章および他の同盟法を考慮した上で、そのような使用の法的根拠となる国内法の範囲内で、そのようなシステムを使用する必要性と比例性を評価しなければならない。その結果、そのようなシステムは、1）FRIA を経た後、2）管轄の国内官庁がその使用を認可した場合にのみ、使用することができる。

(382) リアルタイム RBI システムの使用が認可されるためには、当該加盟国で採択された当該使用を認可する国内法が存在しなければならない。[217]加盟国の中には、データ保護法などの他の連合法または国内法に基づき、生体データシステムの使用を事前に承認する制度をすでに導入している国もある。

10.2.2. 大原則司法当局または独立行政当局による事前認可

(383) AI 法第 5 条(1)(h)(i)から(iii)に列挙された目的のいずれかを追求するリアルタイム RBI システムの使用で、当該加盟国の国内法に規定されているものは、**その使用に先立ち**、司法当局または独立行政当局の認可を受けなければならない。これが大原則である。

(384) ただし、緊急の場合は例外とする。**これには、正当な理由がなければならない[218] 。**緊急**性とは、**「当該システムを利用する必要性が、AI システムの利用を**開始する前に認可を得ることが効果的かつ客観的に不可能である**ような状況」[219] と説明されている。このような緊急性の場合、「AI システムの利用は**必要最小限に**制限されるべきであり、国内法で決定され、法執行機関自身によって個々の緊急性の高い利用ケースに即して指定されるような、**適切な保護措置と条件に従うべきである**」。

10.2.2.1 国内手続規則に従った事前かつ理由ある要請

a) 誰による要請か。

(385) 明記はされていないが、要請は通常、展開者、すなわち**所轄（法執行）当局によって**開始されるものと想定される。AI 法第 3 条(45)b)の法執行機関の定義によれば、「犯罪の予防、捜査、摘発もしくは訴追、または刑事罰の執行（公共の安全に対する脅威の保護および予防を含む）を目的として、公権力および公権力の行使を加盟国の法律によって委託されたその他の団体または事業体」は法執行機関とみなされ、事前認可の要請を提出する「所轄官庁」としての責任を負う可能性もある。

[217]AI 法第 5 条 2 項も参照のこと：AI 法第 5 条 2 項を参照のこと。(...)'.

[218]つまり、「このような場合、法執行機関は、過度の遅滞なく、遅くとも 24 時間以内に、認可を求めることができなかった理由をプロバイダに提供しながら、認可を求めるべきである」ということである。(AI 法前文 35）。

[219]前文 35 AI 法。

(386) AI 法の範囲外の活動にリアルタイム RBI システムを使用する場合、AI 法第 5 条(3)に基づく認可は必要ない。その後、そのようなシステムが法執行目的で使用される場合、その使用は AI 法の範囲に入り、AI 法第 5 条(1)(h)の要件を満たす場合には認可が必要となる。

b) どの用途での要求か？

(387) たとえそのシステムが、スポーツクラブやショッピングモールなど、法執行当局に代わって他者によって運営されている場合であっても**、法執行目的のために、**一般にアクセス可能なスペースで「リアルタイム」RBI システムを使用するには、事前の認可が必要である**。**

例えば、

- 行方不明の子どもを捜索するためのリソースを委託された組織が、リアルタイム RBI システムの使用を決定する。その組織には、公権力や公権力の行使、犯罪の防止、公共の安全に対する脅威の防止といった任務はない。このような利用は、法執行目的ではないため、AI 法第 5 条(1)（h）の禁止事項には該当しない。しかし、そのシステムは、〔附属書 III の 1(a)）「高リスク」に分類され、GDPR 第 36 条に従い、監督データ保護当局の事前協議が必要となる可能性がある。適用法および GDPR 第 9 条 1 項の例外のいずれかが適用されるかどうかによっては、そのような処理に事前の承認が必要となる場合もある。対照的に、同じ組織が法執行当局から、法執行の文脈で、管轄の法執行当局の監督および指示の下で、行方不明の子どもの捜索を代行するよう要請された場合、AI 法第 5 条第 3 項に従って事前の認可が必要となる。

- 自然災害の犠牲者となるリスクのある人々を支援するためのリソースのプロバイダを委託され た民間組織（[220] ）が、その目的のためにリアルタイムの RBI システムを使用することを決定する。このような利用は、法執行目的ではないため、AI 法第 5 条(1)(h)の禁止事項には該当しない。しかし、そのシステムは〔「〕高リスク」（附属書 III の 1（a））に認可され、GDPR 第 36 条に従って監督データ保護当局の事前協議が必要となる可能性がある。適用法および GDPR 第 9 条 1 項の例外のいずれかが適用されるかどうかによっては、そのような処理に事前の承認が必要となる場合もある。

c) いつ？

(388) AI 法第 5 条第 3 項に従い、事前承認は「各使用」に対して必要である。 このことは、このような認可を得る決定的な瞬間は、リアルタイム RBI システムを設置する前の瞬間ではなく、その具体的な使用毎であることを意味する。

例えば、

- 警察がある都市の主要駅にバイオメトリクス対応の CCTV カメラを設置する場合（AI 法に基づく認可は必要ないが、バイオメトリクス・システムは高リスク・システムに関する要件に適合していなければな

[220]自然災害には、河川の氾濫や火災などが含まれる

らず、最初の使用前に FRIA を作成し、システムの個々の使用前に司法当局または独立行政当局による個別の認可が必要である）。

警察は、テロリストが列車で町に到着するという具体的な兆候をつかんでいる（リアルタイムの特定には事前の許可が必要）。

***d)* やる気のあるリクエスト**

(389) AI 法第 5 条 3 項では、リアルタイム RBI の使用に関する個々の要請は、「理由がある」こと、つまり立証され動機づけられたものであることを要求している。

(390) 一部の加盟国は、このような要請をオンラインで提出することを認めている。[221]AI 法第 5 条第 5 項に従い、国内法は、リアルタイム RBI の使用の厳密な必要性と比例性を判断するための十分な証拠、およびそのような使用を許可することの例外的性質を反映するためのその他の関連する側面など、上記の要件を十分に考慮しつつ、 、要請の正確な内容に関する要件を定めるべきである。

10.2.2.2 司法当局または独立行政当局による認可

(391) 認可は、決定を拘束する司法当局または独立行政当局によってのみ認められる。

***a)* 独立認可機関**

(392) CJEU は「独立性」の概念をさまざまな文脈で解釈してきた。例えば、*HK v Prokuratuur* において、CJEU は、認可の独立性とは、当局が「中立的な立場」を維持することを意味すると説明した[222] 。CJEU は、過去の捜査に関与した当局（この場合は検察官）にはそのような独立性はないと明言した。AI 法第 5 条第 3 項が要求する独立性に関しても同様の考慮が適用される可能性があり、これは認可機関が RBI 制度を利用する当局から独立している必要があることを意味する。このことは、警察だけでなく、警察の業務と認可を求める RBI の使用を監督する捜査判事や検察官の場合にも当てはまる。

(393) 欧州委員会対ポーランドの事件では、鉄道の安全という観点から、ある団体がいつ独立しているとみなされるかという問題を扱ったものであるが、CJEU は、「公的機関に関しては、独立性とは通常、当該団体が、その独立性が確保されるべき団体との関係において、いかなる指示や圧力からも遮断され、完全に自由に行動できることを保証する地位を指す」と判断している。[223]同様の指摘は、AI 法第 5 条第 3 項にも当てはまる。

(394) 民主主義社会における司法当局も、一般的には独立した当局である。司法が重要な役割を果たすのは、行政府や立法府から独立し、立法や基本的権利・自由の適用を自律的・独立的な方法で監

[221]例えば、フランスのデータ保護当局である CNIL への認可要請を参照のこと。

[222]2021 年 3 月 2 日の司法裁判所判決、*Prokuratuur*, C-746/18, ECLI:EU:C:2021:152, パラグラフ 54。

[223]2018 年 6 月 13 日の司法裁判所の判決、*Commission* v *Poland*, C-530/16, ECLI:EU:C:2018:430, paragraph 67.

視・審査する場合である。司法の独立は法の支配の重要な側面の一つであり、ECHR 第 47 条（憲章）および第 6 条 1 項によって保証されている。[224]

***b)* 使用が行われる場所の認可**

(395) 認可は、国内法に従って権限を有する当局に宛てて行わなければならない。[225]

***c)* 例外に規定された目的のいずれかを達成するために「必要かつ比例的」な場合にのみ許可する。**

(396) 法執行の目的で、公共のアクセス可能な空間でリアルタイム RBI を使用する認可は、AI 法第 5 条 3 項の要件が満たされているかどうかを評価しなければならない。

侵入性が高い

(397) データ保護の観点から、生体データ、特に顔認識技術の使用は、欧州データ保護委員会（EDPB）のガイドライン 5/2022 および欧州データ保護監察機関（EDPS）により、いくつかの基本的権利と自由に影響を及ぼすとみなされている。この見解は、EU 基本権機関や欧州評議会でも共有されている。[226]欧州司法裁判所（CJEU）[227] と欧州人権高等弁務官事務所（ECtHR）[228] は、生体データ処理の機密性を確認している。

(398) **基本的権利と自由に対する**いかなる干渉も、**常に権利と自由の本質を尊重しなければならない。**これは、憲章第 52 条 1 項からも導かれる。

(399) 基本的権利と自由の「本質」という概念は、欧州司法裁判所の判例において発展してきたものであり、EU の法秩序における独立した価値である。基本的権利や自由の本質が尊重されない場合、それは権利や自由がある措置によって不当に侵害されることを意味するため、前もって干渉することは許されない。

必要かつ適切な場合」に限る

(400) 基本的権利と自由に対するいかなる干渉も、原則として、憲章第 52 条に基づく必要性と比例性を尊重すべき「法律」を必要とする。(AI 法 5 条 5 項以下参照）。AI 法 5 条 3 項は、法執行目的のために公にアクセス可能な空間におけるリアルタイムの RBI の使用を許可する国内法が、「AI 法 5 条 1 項(h)に規定された目的の一つを達成するために必要であり、かつ、それに比例すると（当局が）納得する場合にのみ」、そのような使用の認可が許可されることを規定しなければならないと定めている。

---

[224]R. Manko, , Briefing, European Parliamentary Research Service (EPRS), 2022, 12 p. ; X, *ECJ case law on judicial independence.A Chronological overview*, Briefing, European Parliamentary Research Service (EPRS), 2023, p.12.

[225]2015 年 10 月 6 日の司法裁判所判決（*Schrems*, C-362/14, https://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:62014CJ0362, ECLI:EU:C:2015:650,パラグラフ 44）参照。

[226]個人データの自動処理に関する個人保護条約諮問委員会（ETS 108）、顔認識に関するガイドライン、2021 年。

[227]2023 年 1 月 26 日の司法裁判所の判決、*Ministerstvo na vatreshnite raboti*, C-205/21, ECLI:EU:C:2023:49, paragraphs 60 to 76 and 116 to 134.

[228]2023 年 7 月 4 日の欧州人権裁判所の判決、*Glukhin* v *Russia*, Application no. 11519/20, ECLI:CE:ECHR:2023:0704JUD001151920, paragraphs 88 and 90（以下、「*Glukhin* v *Russia* 判決」という。）

国家当局は、バイオメトリクス識別が厳密に必要かどうかを検証する必要がある。[229]このアセスメントは、具体的な状況を伴う各使用の認可を要求する前に、一般的な 方法で の必要性と比例性のアセスメントをすでに含んでいるはずの FRIA に基づくべきである。

10.2.2.3.事前承認要件の例外：24 時間以内の要請と拒否された場合の結果

(401) 緊急の場合、利用者は、リアルタイム RBI システムが使用された時点から 24 時間以内に承認要請を提出することができる。実際には、これは一般に、バイオメトリクスの準備が整った、または機能を備えたカメラが「ス イッチオン」されて展開され、システムとの最初のバイオメトリクス比較が行われた瞬間となる。要求の適時性を立証するために、処理活動のログを認可が利用できるようにすべきである。 [230]

(402) このような場合、システムを利用する前に、なぜ事前の申請がなかったのか、その理由を説明する必要がある。

10.2.2.4. 承認要求が拒否された場合は直ちに停止し、削除する。

(403) AI 法第 5 条第 3 項はさらに、緊急の場合の認可申請が却下された場合、リアルタイム RBI システムの利用は直ちに中止されるべきであると定めている。このような場合、その使用の結果および出力を含むすべてのデータは直ちに廃棄・削除されなければならない。[231]AI 法第 5 条第 3 項は、この点に関して例外なく明示している。展開者は持つことになる：

a) バイオメトリクス情報（顔画像、音声スニペットなど）および関連する識別情報（該当する場合）を含む参照データベース。

b) 公共のアクセス可能な空間に存在する個人から取得した生体情報を比較し、それらの個人を識別し、特定する。

c) この比較が比較結果につながる。

(404) 収集・処理されたデータを破棄・削除するという要件は、不正な生体データ識別のために使用された参照データベースが、争われた捜査のために特別に構築されたものであれば、それを削除・抹消しなければならないことも意味する。法執行当局が、リアルタイム RBI の不正使用以外の正当な目的のために、合法的な方法で識別に使用されたデータベースを構築し、維持することを意図していた場合に限り、データベースを維持することができる。

---

[229]データ収集についてはこちらも参照のこと：2024 年 11 月 28 日付司法裁判所判決、*Ministerstvo na vatreshnite raboti*, C-80/23, ECLI:EU:C:2024:991.

[230]自動生成的データログは，リスクの高い AI システムについては，少なくとも 6 か月間保存されなければならず，附属書 III の 1(a)に記載されたリスクの高いシステムについては，各使用について，開始データ及び終了データ並びに時刻を含まなければならない。第 12.3 条(a)及び AI 法第 19 条を参照のこと。

[231]監督当局もまた、この事後チェックと管理を行う権限を持つべきである。AI 法第 5 条(5)参照。

(405) 生体情報を含む（違法な）データベースの削除に加え、リアルタイム RBI システムの違法な使用中に取得された、メタデータを含む収集画像およびその他の個人データ、テンプレートを含む技術的処理データ、その他の比較-および 出力データもすべて削除されなければならない。

(406) 法執行当局が拒絶に異議を唱える場合、当該データは、その要請について最終決定がなされるまで、受託者が保管することができる。その期間中、これらのデータは通常、法執行当局の裁量に委ねられるべきではない [310]。

10.2.2.5 リアルタイム RBI システムの出力のみで意思決定を行わない。

(407) AI 法第 5 条第 3 項に従い、リアルタイム RBI システムの展開者が認可を取得した場合でも、「リアルタイム」RBI システムの出力のみに基づいて、人に不利な法的効果をもたらす決定を下すことはできない。

例えば、

- 顔認識システムによる本人確認だけで、それ以上の確認なしに重大犯罪で逮捕・収監される。これは、AI 法第 14 条に規定されている、人間による監視の必要性の上に成り立っている。例えば、ある人物が別の場所にいたことがあるかどうか、あるいはその人物が捜索された人物であるはずがないという他の理由があるかどうか、などである。

AI 法第 *14* 条の人間監視の要件

(408) 第 5 条(1)(h)に列挙された目的のいずれかを追求し、かつ AI 法第 5 条第 2 項～第 6 項を遵守しているために認められているリアルタイム RBI の使用は、依然として高リスクシステムに関する規則に該当する。AI 法第 14 条に従い、リスクの高い AI システムは、「適切なヒューマン・マシン・インターフェース・ツールを含め、それが使用されている期間中、自然人が効果的に監督できるような方法で設計・開発されなければならない」とされている。AI 法第 14 条 5 項に従い、「その識別が、必要な能力、訓練および権限を有する少なくとも 2 人の自然人によって個別に検証・確認された場合」、または「連邦法もしくは国内法がこの要件の適用を不釣り合いとみなす場合」を除き、システムから得られる識別に基づいて、展開者が行動や決定を下してはならない。AI 法第 4 条は、AI システムのプロバイダおよび利用者に対し、「AI システムの操作および利用に携わる職員およびその他の者の十分なレベルの AI リテラシー」を確保し、システムが利用される対象者を考慮するための AI リテラシー対策を規定している。

(409) データ保護の観点から EDPB が述べているように、人間の監視が効果的であるためには、「（この場合は顔認識）システムとその限界を理解し、その結果を適切に解釈できるようにする」ことが重要である。 また、自動化のバイアスの影響を打ち消すような職場と組織を確立し、結果の無批判な受け入れを助長しないようにすることも必要である。例えば、 、時間的なプレッシャー、負担の大きい手続き、

潜在的なキャリアへの悪影響などである□[232]□同様の配慮が AI 法の文脈でも適用される可能性がある

10.3. 法執行のために公共のアクセス可能な空間で「リアルタイム」遠隔バイオメトリクス識別システ ムを使用する場合は、その都度認可に通知する。

> **AI 法第 5 条 4 項はプロバイダを規定している：**
>
> 第 3 項を損なうことなく、法執行目的のために公にアクセス可能な空間における「リアルタイム」遠隔生体デー タ識別システムの各使用は、第 5 項で言及される国内規則に従って、関連する市場監視機関および国 内データ保護機関に届け出なければならない。この届出には、最低限、第 6 項に規定される情報が含まれるものとし、機微な業務データを含んではならない。

(410) AI 法第 5 条(1)(h)(i)(iii)に列挙された目的の 1 つを追求する RBI システムの各使用は、関連する市場監視機関および国内データ保護機関に届け出なければならない。認可の回数とその結果を報告できるようにするため、使用の都度届け出る必要がある。この届出には、機密性の高い業務データを含める必要はない。AI 法第 3 条（38）によれば、機微な業務データ」とは、法執行活動（刑事犯罪の防止、検知、捜査または訴追）に関連する業務データを意味し、その開示は刑事手続きの完全性を危うくする可能性がある。

(411) 報告義務の詳細については、下記 10.6 項を参照のこと。

10.4. AI 法の例外の範囲内での国内法の必要性

10.4.1. 原則：例外のすべてまたは一部について、認可の法的根拠となる国内法が必要である。

(412) 法執行の目的で、公共のアクセス可能な空間における「リアルタイム」RBI システムの使用を運用するためには、国内法が必要である。同時に、AI 法第 5 条(5)は、加盟国がそのような国内法を採択するかどうかを自由に決定できることをプロバイダに定めている。リアルタイム RBI の使用を許可する国内法が採択された場合、AI 法は、国内法が AI 法に規定された要件に準拠するために含まなければならない実質的要素を規定する。

> **AI 法第 5 条(5)**
>
> 加盟国は、(h)(i), (ii), (iii)に列挙された制限および条件の範囲内において、法執行の目的で、公にアクセス可能な空間における「リアルタイム」遠隔バイオメトリクス識別システムの使用を全面的または部分的に認可する可能性を規定することを決定することができる。関係加盟国は、第 3 項で言及される認可の要求、発行および行使、ならびに監督および報告に関する必要な詳細規則を国内法に定めるものとする。これらの規則はまた、第 1 項 (h)のどの目的 ((h)(iii)に言及される犯罪のどれを含

[232]EDPB, *Guidelines 05/2022 on use of facial recognition technology in area of law enforcement* Version 2.0, 26 April 2023, p. 22.

む）に関して、権限のある当局が法執行の目的でこれらのシステムを使用する認可を受けることができるかを規定するものとする。加盟国は、遅くとも採択後 30 日以内に、これらの規則を欧州委員会に通知しなければならない。加盟国は、連邦法に従い、遠隔生体認証システムの使用に関して、より制限的な法律を導入することができる。

10.4.2. 国内法は、AI 法第 5 条(1)(h)の制限と条件を尊重するものとする。

(413) 法執行を目的とした、公共のアクセス可能な空間における「リアルタイム」RBI システムの使用は、基本的権利の侵害とみなされるため、AI 法第 5 条(5)は、このような使用は加盟国の国内法によって確立されるものと定めている。これらの国内法は、このようなシステム使用の法的根拠を提供する。

(414) 国内法は、AI 法第 5 条(1)(h)が定める制限を超えてはならず、また AI 法に規定されるすべての関連条件を尊重しなければならない。つまり、加盟国は、法執行のために公共のアクセス可能な空間でリアルタイム RBI を使用できる目的を、AI 法第 5 条(1)(h)(i)～(iii) [233] に列挙されている目的以上に拡大してはならない。

(415) 加盟国は、国内法を採択した後、遅くとも 30 日以内に欧州委員会に通知しなければならない。この通知は、加盟国の法律が AI 法に適合していることを推定するものではない。 AI 事務局は通知を受領した後、受領確認を送付する。採択に先立ち、加盟国は提案された国内法（または地域法）の暫定版を AI 事務局に送付することも奨励される。いずれにせよ、第 5 条(5)に規定された採択後 30 日以内という法的期限内に AI 事務局への通知が行われなかった場合、その国内法は法的手続において執行不能となる可能性がある。[234]欧州委員会は加盟国の法律を公開ウェブサイトで公表する。

(416) 加盟国は、連合法に従い、より制限的な法律、すなわち AI 法第 5 条(1)(h)および(2)～(7)に規定された要件よりも厳しい要件を持つ法律を導入することができる。

10.4.3. 認可申請、発行、行使に関する国内法の詳細

(417) 認可の要請、発行、および行使に適用される詳細な規則に関しては、国内法によって定め られる。この規則は、リアルタイム RBI システムの使用に関する適切かつ完全な情報を認可当局に提供し、当該使用の厳密な必要性および比例性について認可当局が判断できるようにすることを目的としている。

リアルタイム RBI システムの使用を許可する国内法は、例えば以下のような規制を設けることができる、

- AI 法第 5 条(1)(h)の管轄当局であり、認可を発行（または認可を拒否）する権限を有する加盟国の独立当局である；

[233]2022 年 4 月 5 日司法裁判所判決、*An Garda Síochána* 委員、C-140/20、ECLI:EU:C:2022:258、パラグラフ 54 参照：「比例性の要件を満たすためには、国内法は当該措置の範囲と適用を規定し、最低限のセーフガードを課す明確かつ正確な規則を定めなければならない。

[234]類推により、2019 年 12 月 19 日の司法裁判所の判決、*Airbnb Ireland*, C-390/18,EU:C:2019:1112, paragraphs 96 to 97 を参照のこと。

- 公にアクセス可能な空間におけるリアルタイム RBI を法執行目的に使用することができる目的の詳細な範囲（第 5 条(1)(h)(i)～(iii)に列挙された目的を超えることなく、場合によってはさらに絞り込む；

- プロバイダは、要請を文書で行い、その使用を正当化する具体的な犯罪や状況について、具体的な使用方法と使用目的の詳細な説明を求める；

- AI 法第 5 条(1)(h)(i)～(iii)に列挙された目的、特に場所、期間、個人の範囲に関連し、システムの使用の関連性、十分性、効率性、より侵入的でない手段がないことを含め、厳格な必要性と比例性を正当化するシステムの使用を正当化するための動機付けと裏付け証拠の提出（および関連する場合は翻訳の必要性）が要求される；

- 使用される技術の説明とデータ収集の場所；

- 使用されるシステムの最低信頼性、使用されるしきい値、および精度率；

- 技術的な詳細や規準の正確さを含め、提出された情報を認可機関が事前・事後にいつでも監査できること；

- 使用する参照データベースの指定；

- 収集されたデータおよび使用されたその他すべての関連個人データの保存期間；

- データへの不正アクセス対策を含むセキュリティ対策；

- その他のセーフガード（関連する場合）；

- 他国を含む、民間または公的機関との協力、データの移転および交換に関する記述；

- プロセスのトレーサビリティである；

- 展開責任者の名前；

その他の形式的

- 審問を補完する書面手続の可能性；

- 拒否の理由

- 捜査の対象となる人物の権利、個人データを取得された人物の権利、第三者の権利 ；[235]

- 認可当局が決定を下すまでの時間である；

[235]例えば、EDPB, *Guidelines 05/2022 on use of facial recognition technology in area of law enforcement* Version 2.0, 26 April 2023, p. 24 et seq.

- 認可の付与／拒否に際しての正式な通知の必要性；

- 正式かつ実質的な）要件に従わない場合は制裁を科す；

- 拒否された認可を不服とする権利；

練習として

- リアルタイム RBI システムの使用を中央登録簿に登録し、実質的な要素を要約すること；

- さらなる報告義務の可能性もある；

- 認可を延長または変更するための手続き。

10.4.4. 認可に関する監督と報告に関する詳細な国内法

(418) AI 法第 70 条は、加盟国に『少なくとも 1 つの認定機関と 1 つの市場監視機関を設置する』ことを義務付けている。AI 法第 74 条 8 項は、『加盟国は、規則（EU）2016/679 もしくは指令（EU）2016/680 に基づく管轄データ保護監督当局、または指令（EU）2016/680 の第 41 条から第 44 条までに規定された同一の条件に従って指定されたその他の当局のいずれかを、本規則の目的上、市場監視当局として指定しなければならない』と規定している。

(419) これは、加盟国が AI 法第 5 条(1)(h)(i)から(iii)に列挙された目的のためにリアルタイム RBI システムの使用を認可する前に設けなければならない認可当局の指定に加えて行われる。

10.5. 加盟国の市場監視当局およびデータ保護当局による年次報告書

AI 法第 5 条 6 項は次のようにプロバイダを規定している。

各国の市場監視当局および各国のデータ保護当局は、以下の通りである。

第 4 項に従い、法執行の目的で、公にアクセス可能な空間における「リアルタイム」遠隔バイオメトリクス識別システムの使用を通告された加盟国は、当該使用に関する年次報告書を欧州委員会に提出しなければならない。そのために、欧州委員会は、加盟国および各国の市場監視およびデータ保護当局に対し、管轄司法当局または独立行政当局が第 3 項に従って認可の要請に対して下した決定の数およびその結果に関する情報を含むテンプレートを提供する。

(420) 公共のアクセス空間におけるリアルタイム RBI システムの法執行目的での使用（第 5 条(4)参照）について、展開者から報告を受けた加盟国の市場監視当局およびデータ保護当局は、そのような使用に関する年次報告書を欧州委員会に提出しなければならない。これらの報告は、欧州委員会が提供するテンプレートに基づいて行われる。この雛形はいずれ作成される予定である。

(421) 展開主体が EU の機構、団体、機関である場合、EDPS は、法執行のために一般にアクセス可能な空間で使用されているリアルタイム RBI システムについて、年 1 回、欧州委員会に報告する義務を負う。

(422) AI 法は加盟国に対し、2025 年 2 月 2 日から 2025 年 8 月 2 日までの期間を対象としているため、国内データ保護当局の報告書のみが対象となる。

(423) 各国の市場監視当局とデータ保護当局は、加盟国ごとに個別の報告書を提出するか、共同報告書を提出するかを自由に決めることができる。

10.6. 委員会による年次報告

AI 法第 5 条 7 項は次のようにプロバイダを規定している。

欧州委員会は、第 6 項の年次報告書に基づいて加盟国で集計されたデータに基づき、法執行を目的とした、公共のアクセス可能な空間におけるリアルタイムの遠隔生体認証システムの使用に関する年次報告書を発行する。これらの年次報告書には、関連する法執行活動の機微な運用データは含まれないものとする

(424) AI 法は、欧州委員会に対し、 加盟国および欧州連合の機構、機関、団体による、法執行を目的とした、一般にアクセス可能な空間におけるリアルタイム RBI システムの利用について、集計データに基づく年次報告書を発行することを義務付けている。これらの報告書は、AI 法第 5 条第 6 項に従って各国認定機関から認可された情報に基づいて作成される。

(425) 欧州委員会の年次報告書には、機微な業務データは含まれてはならない。機微な業務データとは、「刑事犯罪の防止、検知、捜査または訴追の活動に関連する業務データであって、その開示が刑事手続きの完全性を危うくする可能性のあるもの」を意味する。[236]これは、例えば場所や使用されたカメラなど、進行中または過去の捜査を明らかにする具体的な詳細を公表してはならないことを意味する。

10.7. 対象外

(426) AI 法第 5 条(1)(h)の禁止事項の対象外である RBI システムのその他の使用はすべて、AI 法の範囲内であれば、第 6 条で定義され、AI 法附属書Ⅲの 1(a)に記載されている高リスク AI システムのカテゴリーに含まれる。

(427) AI 法第 5 条(1)項(h)の禁止範囲から外れる RBI システムには、バイオメトリクス検証／認証システムや、法執行を目的とした、一般にアクセス可能な空間における（事後）RBI システムの遡及 的利用が含まれる。例えば、警察当局は、犯罪容疑者の画像を犯罪データベースの記録された顔画像と比較するために、遡及的な顔認識を行うことを国内法によって認可される場合がある。[237]また、私的

[236]AI 法第 3 条第 38 項

[237]例えば、フランスの *Traitement des Antécédents Judiciaires* データベースは、*Décret no. 2012-652 du 4 mai 2012 relatif au Traitement des Antécédents Judiciaires*（法令 2012-652）により作成された。

な空間（誰かの家など）またはオンライン空間（児童性的虐待資料を流布した容疑者を特定するためのチャットルームやオンラインゲームの利用など）での法執行を目的としたリアルタイムの RBI システムの利用も、禁止の範囲外の利用である。最後に、民間の主体による RBI システムの使用は、リアルタイムでも遡及的でも（例えば、スーパーマーケットによる、既知の万引き犯を特定するためのライブ顔認識技術の使用、スポーツ競技場による、競技場への入場が禁止された個人を特定するためのライブ顔認識技術の使用、あるいは学校における、セキュリティ目的や登校のためのライブ顔認識技術の使用など）、禁止の範囲外となる。

(428) 一般的にリスクの高い AI システムに適用される規則に加えて、法執行目的での **RBI システムの遡及的使用**は、AI 法第 26 条 10 項（2026 年 8 月 2 日から適用）に従い、追加条件と保護措置の対象となる。[238]。

(429) **法執行以外の目的での**使用は、いかなる場合でも**データ保護規則を**遵守しなければならない。以下のケースは、そのような使用のケースにおける GDPR 第 9 条 2 項の解釈と、生体データを処理するための例外を示している。

例えば、

- フランスの行政裁判所は、2 つの公立学校でアクセス管理およびセキュリティ目的でライブ顔認識技術を試用したことは、（データ保護規則上）必要性も比例性もないと判断した。バッジの使用など、生徒にとってより侵害の少ない代替案が利用可能であった。さらに、明示的な同意の条件を満たしていなかった。したがって、高校で顔認識技術を試用する妥当な法的根拠として同意を用いることはできなかった。[239]

- オランダで、あるスーパーマーケットが万引き防止のためにライブの顔認識技術を使用することは許されなかった。顧客からの明示的な同意や、実質的な公共の利益（セキュリティ目的など）のための処理を認める法的根拠がなければ、スーパーマーケットは生体データを処理することができず、その結果、顔認識技術を展開することができなかった。[240]

- サポーターを識別するためにサッカークラブの入口でライブの顔認識技術を使用することはフランスで禁止され（[241]）、観客の安全を確保するために使用することはスペインで禁止された。[242]

[238]AI 法第 26 条 10 項および前文 94 項。

[239]TA マルセイユ（マルセイユ行政裁判所）2020 年 2 月 27 日、第 1901249 号。

[240]https://www.autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-issues-formal-warning-to-supermarket-for-use-of-facial-recognitiontechnology。

[241]https://www.cnil.fr/fr/reconnaissance-faciale-et-interdiction-commerciale-de-stade-la-cnil-adresse-un-avertissement-un-club

[242]https://www.biometricupdate.com/202401/spanish-data-authority-opposes-facial-recognition-for-football-stadium-access

10.8. 使用例

警察は、欧州選手権の試合中、サッカースタジアムの正面玄関周辺の警察車両に AI ベースの顔認識技術を搭載したモバイル CCTV カメラを設置し、周辺の安全を確保するとともに、指名手配者のアドホック・ウォッチリスト・データベースに顔が記録されている人物を特定する。この監視リストには、犯罪を犯した疑いのある人物（重大犯罪から詐欺、強盗に至るまで）、諜報目的で関心を持つ可能性のある人物、精神的問題を抱えた脆弱性のある人物などが含まれている。警察のライブ顔認識技術の使用は、イベントに特定の人物がいるかどうかという情報とはリンクしていない。リアルタイムの RBI の使用が許可される人物は、捜索のための監視リストに載っている可能性が高いが、このリストはあまりにも具体性に欠け、サッカーの試合というイベントとリンクしていない。 したがって、このような使用は禁止されている。

バイオメトリクス認証システム（遠隔操作ではない）が、原子力発電所への立ち入りの可否を確認する。人々が（明らかな）カメラの前に姿を現し、システムによってアクセスが拒否されると、システムはその後、その人物がテロリストの監視リストに載っているかどうかを識別しようとする。システムは遠隔操作ではない。原発への入場を許可するために、人々は積極的に検証に参加していた。

この使用例は、AI 法第 5 条の禁止事項には該当しない。

多忙な都市の警察当局は、ライブ顔認識技術を実行できる AI 搭載の CCTV カメラを展開している。おそらく、顔認識の上に、物体検知や群衆の動きなど、さまざまな機能が追加されているのだろう。

彼らはこれらのカメラを、礼拝所、LGBT+コミュニティがよく訪れる多くの場所、医院、薬局、さまざまなレストランやバーなど、複数の場所に設置している。

このような生体認証対応カメラの設置は、AI 法では禁止されていない。

ただし、不特定かつ無差別な自然人の特定など、特定の利用は禁止されている。

夏休み中、ある住宅街で数件の強盗事件が発生した。警察は、強盗が発生する前に近隣で容疑者を目撃した目撃者から容疑者の特徴を聞き出す。容疑者を特定し逮捕するため、警察は週末に近所のさまざまな場所でライブ顔認識技術を使用する。目撃者の指示に基づいて、警察は容疑者の顔合成を作成し、保管データベースから顔合成に似た人物の写真を数枚抽出した。

たとえ警察が、対象となる容疑者に対してライブの顔認識技術を使用し、使用範囲と使用時間を定めたとしても、AI 法の附属書 II に記載されていない犯罪の場合には、その使用は展開できない。

警察は、サッカースタジアムにいるファンの感情を生体認証システムでスクリーニングしている。このシステムは、潜在的な攻撃性を発見すると、直ちにスタジアムのその部分にリアルタイムの RBI を展開し、過去に暴力を振るったフーリガンを特定する。

スタジアムでの感情の選別は、AI 法では禁止されていない（それでも AI 法の高リスクカテゴリーに該当する）。しかし、リアルタイム RBI の適用は、特に、法執行の目的で人物を特定する必要性を決定するのがバイオメトリクス・システムである場合、AI 法で禁止される。

警察は、街頭で集団抗議行動を組織した政治的抗議者を特定するために、市内や地下鉄に設置された CCTV ネットワークに依存している。当該加盟国では、公道や道路などの公共の場所で行われる集団的抗議活動の認可者は、公共の混乱や暴力を防ぐため、抗議活動の 3 日前に市当局に届け出なければならない。届け出がない場合は、6 カ月以下の禁固刑と最高 8,000 ユーロの罰金が科される犯罪行為となる。デモ参加者を特定するため、警察は街頭に設置された CCTV カメラの映像を抽出し、抽出された画像とソーシャルメディアに投稿された写真を比較することで、遡及的に顔認識を行う。

顔認識技術の遡及的使用は AI 法で禁止されていない。その使用はリスクが高いと考えられ、そのようなシステムに関する AI 法の要件に従うべきである 。[243]

禁止されていない行為の例

- ホテルはリアルタイム RBI を使って VIP 客を認識する。これは法の執行ではない。
- ショッピングモールがリアルタイムの RBI を使って万引き犯を見つける。これは法の執行ではない。

禁止されている：

警察から委託を受けたショッピングモールは、万引き犯を見つけるためにリアルタイムの RBI を使用している。このシステムは法執行の目的で、公共のアクセス可能な場所に展開されている。万引き犯の捜索は AI 法第 5 条(1)（h）のどの例外にも該当しないため、使用は禁止されている。

[243]法執行を目的とした生体データの処理は、依然として LED 第 10 条の対象であり、これは国家レベルで実施される必要がある。FRT の遡及的使用を行うための処理は、厳密に必要な場合にのみ許可されるべきであり、適切なセーフガードの対象となるべきである。FRT の遡及的利用が、デモ参加者を特定するために厳密に必要かどうかは疑問である。このシナリオの根拠となるグルヒン対ロシア判決において、ECtHR は、犯罪検知は正当な目的でありうるが、公共秩序や交通の安全に対するリスクがない以上、FRT の使用は遡及的であれライブであれ、不釣り合いであると裁定した。裁判所は、FRT の「非常に侵入的」な性質を強調した。裁判所は、FRT の使用は差し迫った社会的必要性に応えるものではなく、民主主義社会において必要なものでもないと結論づけた。

11. 適用

(430) AI 法第 113 条によれば、AI 法第 5 条は 2025 年 2 月 2 日から適用される。同条項の禁止事項は、その日の前後を問わず、原則としてすべての AI システムに適用される[244] 。

(431) 同時に、ガバナンス、エンフォースメント、罰則に関する章は 2025 年 8 月 2 日から適用される。従って、AI 法第 5 条の禁止事項の不遵守に対する罰則規定は、2025 年 8 月 2 日以前には適用されない。この暫定期間には、禁止事項が適切に遵守されているかどうかを監視する市場監視当局も存在しない。

(432) とはいえ、この暫定期間であっても、AI システムのプロバイダや展開事業者には、禁止事項が全面的に適用され、義務付けられる。したがって、これらの事業者は、AI 法第 5 条の禁止行為に該当する可能性のある AI システムを上市、運用、使用しないよう、必要な措置を講じる必要がある。監視や罰金に関する規定が適用されるのが遅くなるとしても、禁止事項そのものは直接的な効力を持つため、影響を受ける当事者は国内の裁判所で強制執行を行い、禁止行為に対する仮処分を請求することができる。

12. 委員会ガイドラインの見直しと更新

(433) 本ガイドラインは、AI 法第 5 条における禁止事項の実践的な事例を用いた最初の解釈である。欧州委員会は、AI システムのプロバイダや展開者、AI 委員会、その他の関係者の意見を聞きながら、事業者や当局が禁止事項をどのように理解し、さらに実用的なユースケースを収集するかを継続的に支援していく。

(434) 欧州委員会は、禁止事項の実施において得られた実務経験や、この分野における技術的、社会的、規制的発展のペースを考慮し、必要に応じ速やかに本ガイドラインを見直す。これには、本ガイドラインで検討されている禁止事項および AI 法のその他の規定に関して、市場監視の強制措置や欧州司法共同体（EU）が下した解釈から得られた関連する経験も含まれる。このような見直しの過程において、欧州委員会は本ガイドラインの撤回または修正を決定することができる。欧州委員会は、AI システムのプロバイダや展開者、AI 委員会を通じた各国の市場監視当局、AI アドバイザリー・フォーラム、研究機関、市民団体に対し、今後の公開協議の呼びかけに応じることで、このプロセスに貢献することを奨励する。

[244]AI 法第 111 条第 1 項および第 2 項を参照のこと。この条文では、祖父条項が、AI 法第 113 条第 3 項（a）で言及されている AI 法第 5 条の適用を損なうものではないことが明記されている。