外国著作権法令集(64)

― EU AI 規則 ガイドライン編 ―

AI 規則５条のガイドライン

井奈波　朋子 訳

March 2026

公益社団法人 著作権情報センター

欧州委員会

ブリュッセル、2025 年 2 月 4 日
C(2025) 884 最終
別紙

委員会通達別紙

委員会からの通達案の内容承認 -
規則(EU)2024/1689（AI 法）に定める人工知能について禁止される行為に関する委員会ガイドライン

目　次

1. 背景および目的

(1) 人工知能に関する統一ルールを定め、かつ一定の規則を改正する 2024 年 6 月 13 日の欧州議会および欧州理事会の規則(EU)2024/1689（「AI 法」）[1]は、2024 年 8 月 1 日に発効した。AI 法は、EU における人工知能（「AI」）の上市、サービス開始および使用に関する統一ルールを定める[2]。その目的は、AI のイノベーションおよび採用を促進すると同時に、民主主義および法の支配を含む、EU における健康、安全および基本的権利の高度な保護を確保することである。

(2) AI 法は、リスクベースアプローチに従い、AI システムを 4 つの異なるリスクカテゴリに分類する。

(i) 容認できないリスク：基本的権利および EU の価値観に容認できないリスクをもたらす AI システムは、AI 法第 5 条に基づき禁止される。

(ii) ハイリスク：健康、安全および基本的権利に高度なリスクをもたらす AI システムは、一連の要件および義務の対象となる。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 条に従い「ハイリスク」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 条に基づき透明性義務の対象となる。

(iv) ミニマルリスクからノーリスク：ミニマルリスクまたはノーリスクである AI システムは規制されないが、提供者および導入者は、自主的な行動規範を自発的に遵守できる。[3]

(3) AI 法第 96 条第 1 項(b)に従い、欧州委員会は、AI 法第 5 条に基づき禁止される行為の実務的な実施に関するガイドラインを採択する予定である。これらの禁止は、AI 法の発効から 6 か月後、すなわち 2025 年 2 月 2 日から適用される。

(4) これらのガイドラインは、その一貫性のある、効果的かつ統一的な適用を確保するために、AI 法第 5 条における禁止事項について、法的安全を高め、かつ欧州委員会の解釈に関する見解を示すことを目的とする。これらは、AI 法に基づく管轄当局の執行行為を支援し、さらに、AI システムの提供者および導入者による AI 法に基づくその義務の遵守を確保するため、実務的ガイダンスとして役に立つものでなければならない。これらは、イノベーションを促進しつつ、かつ法的安全を提供しつつ、基本的権利および安全を保護するために AI 法の目的を達成するふさわしい方法において、禁止事項を解釈するよう努めている。

(5) これらのガイドラインには法的拘束力がない。AI 法のあらゆる正式な解釈は、最終的には欧州司法裁判所（CJEU）によってのみ示され得る。

---

[1] 人工知能に関する統一ルールを定める 2024 年 6 月 13 日の欧州議会および欧州理事会規則(EU)2024/1689(人工知能法)(OJ L、2024/1689、2024 年 7 月 12 日)。

[2] AI 法第 1 条

[3] AI 法第 95 条

(6)　これらのガイドラインの草案は、たとえば、AI システムの提供者および導入者、市民社会組織、学界、公的機関、業界団体など、さまざまな利害関係者から、欧州委員会により組織された広範な協議プロセスを通じて収集された情報をもとに作成された。AI 委員会における加盟国および欧州議会からも意見を得た。これらのガイドラインは、AI 法第 5 条の実務上の実施および技術的な動向や市場の動向から得られる経験に照らし、定期的な見直しが予定されている。

(7)　AI 法第 5 条の適用には、個別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評価が求められることになる。したがって、これらのガイドラインにおいて与えられる例は単なる指標であり、それぞれのケースにおけるそのような評価の必要性を損なうものではない。

2. AI に関して禁止される行為の概要

(8)　AI 法第 5 条は、一定の AI システムを、その固有の性質により、基本的権利および EU の価値観を侵害する操作、搾取、社会統制または監視行為のために EU 市場に上市し、サービスを開始し、または使用することを禁止する。AI 法前文 28 項が明確にするとおり、そのような行為は、特に有害で濫用的であり、かつ人間の尊厳、自由、平等、民主主義、法の支配を尊重する EU の価値観、および、欧州連合基本権憲章(「憲章」)に記されている、差別されない権利（憲章第 21 条)および平等(第 20 条)、ならびに、データ保護(憲章第 8 条)および私生活と家庭生活(憲章第 7 条)、ならびに子どもの権利(憲章第 24 条)を含む、基本的権利に反することから、禁止されなければならない。AI 法第 5 条の禁止はまた、表現および情報の自由(憲章第 11 条)、集会および結社の自由(憲章第 12 条)、思想、良心および宗教の自由(憲章第 10 条)、効果的な救済および公正な裁判を受ける権利（憲章第 47 条)、無罪の推定および弁護権(憲章第 48 条)を支えることも目的とする。

2.1. AI 法第 5 条に列挙される禁止事項

(9)　禁止事項の概要

規定	禁止事項	内容
5 条 1 項 a	有害な操作・欺瞞	人の認識の域を超えたサブリミナル技術、または意図的な操作技術もしくは欺瞞的な技術を用いる AI システムであって、行動を歪曲する目的または効果を有し、重大な害を引き起こしまたは合理的に引き起こし得る AI システム
5 条 1 項 b	有害な脆弱性の悪用	年齢、心身障害、または特定の社会的または経済的状況に起因する脆弱性につけこむ AI システムであって、行動を歪曲する目的または効果を有し、重大な害を引き起こしまたは合理的に引き起こし得る AI システム

5条1項c	ソーシャルスコアリング	社会的行動または人もしくは人の特徴に基づいて自然人または人のグループを評価しまたは分類するAIシステムであって、データが無関係な社会的文脈から得られた場合に当該ソーシャルスコアが有害または不利な取扱いにつながるもの、または当該取扱いが社会的行動に照らして不当または不均衡であるもの
5条1項d	個人の刑事犯罪リスク調査および予測	プロファイリングまたは人格的特徴または特性のみに基づいて、人が刑事犯罪を犯すリスクを評価または予測するAIシステム。ただし、犯罪行為に直接結びつく客観的かつ検証可能な事実に基づき、人による評価をサポートするためである場合を除く
5条1項e	顔認識データベース開発のための無差別なスクレイピング	インターネットやCCTV映像から顔画像を無差別にスクレイピングすることにより、顔認識データベースを作成または発展させるAIシステム
5条1項f	感情認識	職場または教育機関において感情を推測するAIシステム。ただし、医療上または安全上の理由による場合を除く。
5条1項g	生体分類	人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向を推測するために、生体データに基づいて人を分類するAIシステム。ただし、法執行分野を含む、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を除く。
5条1項h	リアルタイム遠隔生体識別（RBI）	法執行目的による公衆がアクセス可能な場所内のリアルタイム遠隔生体識別のためのAIシステム。ただし、特定の被害者を対象とする捜索、テロリスト攻撃を含む特定の脅威の防止、または特定の犯罪の被疑者の捜索に必要な場合を除く（許可を含む更なる手続的要件は、AI法5条2項―7項に概説）。

2.2. 禁止の法的根拠

(10) AI法は、2つの法的根拠によって支えられている：欧州連合機能条約(TFEU)第114条(域内市場法制)とTFEU第16条(データ保護法制)である。TFEU第16条は、法執行目的による遠隔生体識別(RBI)システムの使用禁止、法執行目的による生体分類システムの使用禁止、および法執行目的による個々のリスク評価の使用禁止に関し、個人データ処理に関する特定のルールの法的根拠として機能する[4]。AI法第5条に列挙されるその他のすべての禁止は、TFEU第114条に法的根拠がある。

[4] AI法前文3項。TFEU第16条に基づく禁止事項については、アイルランドとデンマークのための2つの関連するオプトアウトがある。TEUおよびTFEU附属の自由、安全、司法分野（AFSJ）における英国およびアイルランドの地位に関するプ

2.3. 実体的範囲：AI システムの「上市」、「サービス開始」、または「使用」に関する実務

(11) AI 法第 5 条により禁止される行為は、特定の AI システムの上市、サービス開始または使用に関係する[5]。リアルタイム遠隔生体識別(RBI)システムに関しては、AI 法第 5 条第 1 項(h)の禁止は、それらの使用にのみ適用される。AI 法第 3 条(1)は、何が AI システムを構成するかについて定義する。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定める。

(12) AI 法第 3 条（9）によれば、AI システムの**上市**は、「AI システム[...]を、EU 市場において初めて利用に供すること」である。「利用に供する」とは、「有償または無償で、商業活動の範囲内で、EU 市場における頒布または使用のために」システムを供給することと定義される。[6] AI システムを利用に供する行為は、たとえば、アプリケーション・プログラミング・インターフェース(API)、クラウド、直接のダウンロード、物理的なコピー、または物理的な製品への搭載により、システムおよびそのサービスへアクセスするなどの供給手段を問わず対象となる。

たとえば、第三国の提供者により EU 域外で開発された RBI システムは、それが一以上の加盟国において有償または無償で提供される場合、初めて EU 市場に上市される。そのような上市は、API または他のユーザーインターフェースを通じ、システムに対するアクセスをオンラインで提供することにより行われる可能性がある。

(13) AI 法第 3 条（11）は、**サービス開始**を「AI システムの意図目的に従って、EU 域内において導入者に最初に使用させるためまたは自ら使用するために、AI システムを提供すること」と定義しているので、最初の使用のための第三者への提供、ならびに社内での開発および導入の双方を対象とする。システムの意図目的とは、「提供者が、使用説明書、販促資料、販売資料、および指示書ならびに技術文書において伝達する情報に明示するような、特定の使用状況および使用条件を含む、提供者が AI システムの目的とする使用」をいう。[7]

たとえば、提供者が、EU 域外で RBI システムを構築し、そのシステムを加盟国内の法執行機関または民間企業に対し最初に使用させるために提供すると、そこでサービス開始となる。 たとえば、公的機関が、社内でスコアリングシステムを開発し、かつそれを導入し、家計手当受給者の不正リスクを予測すると、そこでサービス開始となる。

(14) AI システムの「**使用**」は AI 法において明確に定義されていないが、上市後またはサービス開始後の、そ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導入を対象とする

ロトコル第 21 号に基づき、アイルランドに与えられた裁量により、アイルランドは、法執行目的による公共の場所における RBI のリアルタイム使用禁止に関する規則、および同条に関連する手続上のルール（AI 法第 5 条第 2 項ないし第 6 項）を適用しないことを決定することができる(前文 40 項参照)。デンマークは、TEU および TFEU 附属のプロトコル第 22 号を適用する場合のオプトアウト合意の利益を享受し、TFEU 第 16 条に基づく禁止を完全には適用しないことを決定することができる(前文 41 項参照)。

[5] これらの用語の定義については、欧州委員会通知 –2022 年 EU 製品ルールの実施に関する「ブルーガイド」、2022/C 247/01、第 2 節も参照。

[6] AI 法第 3 条（10）

[7] AI 法第 3 条（12）

広い意味で理解されなければならない。これには、より複雑なシステム、プロセスまたはインフラの一部を含む、AI システムを利用する人（または複数の人）のサービスおよびプロセスに、AI システムを組み込むことも含まれ得る。AI システム提供者が、その AI システムを上市する前に合理的に予見し得る使用条件を考慮しなければならないとしても(意図された使用および合理的に予見可能な誤用[8])、導入者は、システムの使用について適法要件を受け入れることに対する責任を負い続ける[9]。AI 法第 5 条の目的から、「使用」とは、禁止される行為も同然の AI システムのあらゆる誤用(「合理的に予見可能」か否かを問わない)も含まれると理解されなければならない。[10]

たとえば、雇用者が職場において感情を推測するために使用する AI システムは、医療目的または安全目的で使用される場合を除き、禁止される(AI 法第 5 条 1 項(f))。このような禁止は、提供者(システムの供給者)が、導入者(雇用者)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どうかにかかわらず、導入者に適用される。

2.4. 人的範囲：責任を負う行為者

(15) AI 法は、AI システムに関して、さまざまなカテゴリーのオペレータを区別する：すなわち、提供者、導入者、輸入者、頒布者、製品メーカー。本ガイドラインは、AI 法第 5 条の禁止される行為の範囲を考慮し、提供者および導入者のみに焦点を当てる。

(16) AI 法第 3 条(3)によれば、提供者とは、AI システムを開発しまたは開発させ、かつ自己の名または自己の商標のもとで、それを上市し、またはそのサービスを開始する、自然人または法人、公的機関、事務所、またはその他の組織をいう[11]（上記 2.3 参照)。EU 域外に設立または所在する提供者は、それらのシステムを EU 市場に上市するか、もしくはサービスを開始する場合[12]、または AI システムの出力が EU 域内において使用される場合[13]、AI 法の規定の適用対象となる。提供者は、AI システムを上市しまたはサービスを開始する前に、その AI システムが関連するすべての要件を満たすことを確保しなければならない。

たとえば、RBI システムの提供者は、その商標のもとで EU 域内においてシステムを販売する製造業者である。このようなシステムの提供者は、システムを社内で開発し、かつ自身の使用のためにサービスを開始する公的機関である可能性もある。

[8] AI 法第 3 条（12）および（13）参照。
[9] これらの用語の定義については、欧州委員会通知 －2022 年 EU 製品ルールの実施に関する「ブルーガイド」、2022/C 247/01、第 2.8 節も参照。
[10] AI 法前文 28 項。
[11] AI 法第 3 条（3)、(9）および（11)。ハイリスク AI システムに関し、AI 法第 25 条第 1 項は次のように規定する。すべての頒布者、輸入者、導入者またはその他の第三者は、本規則の目的上、ハイリスク AI システムの提供者とみなされ、以下の状況のいずれかにおいて、第 16 条に基づく提供者に課される義務の対象となる：(a) 別段の義務を定める契約上の規定を害することなく、既に上市されまたはサービスが開始されているハイリスク AI システムを、その名またはその商標のもとで販売する場合 (b) 既に上市されまたはサービスが開始されているハイリスク AI システムに対し、実質的な変更を行い、第 6 条の適用によりそれが引き続きハイリスク AI システムとなる場合；(c) ハイリスクに分類されず、既に上市されまたはサービスが開始されている AI システム（汎用 AI システムを含む）の意図目的を変更し、関係する AI システムが第 6 条に基づきハイリスク AI システムとなる場合。
[12] AI 法第 2 条第 1 項(a)
[13] AI 法第 2 条第 1 項(c)

(17) **導入者**は、その権限の下でAIシステムを使用する自然人または法人、公的機関、事務所、またはその他の組織であり、事業性のない個人的な活動の範囲内における使用を除く[14]。AIシステムに対する「権限」とは、システムを導入することの決定に対し、およびその実際の使用方法に対し、責任を負うことと理解されなければならない。導入者は、その設立場所または所在地がEU域内にある場合[15]、または第三国に所在しているのであればAIシステムの出力がEU域内で使用される場合に、AI法が適用される[16]。

(18) その権限の下でシステムが使用されるAIシステムの導入者が法人、すなわち、法執行機関または民間の警備会社である場合、手続きの範囲内で、かつそれらの者の管理下において行動する個々の従業員は、導入者であるとみなされない。法人は、第三者（たとえば、請負業者、外部スタッフ）が、法人に代わって、かつその責任と監督の下でシステムの運用に関与する場合にも、依然として導入者である。

(19) オペレータは、AIシステムに関し、複数の役割を同時に果たすことができる。たとえば、オペレータが、独自のAIシステムを開発し、その後使用する場合、たとえそのシステムが有償または無償で他の導入者に提供され、その者によっても使用されるとしても、そのシステムの提供者および導入者の双方であるとみなされる。

(20) AI法の継続的遵守は、AIのライフサイクルの全ての段階において求められる。これは、AIシステムがそのライフサイクルを通じてAI法を遵守し、かつAI法第5条に基づき禁止される行為に至らないことを確保するため、EU市場に上市されまたはサービスが開始されたAIシステムの継続的なモニタリングとアップデートを必要とする。AIシステムの提供者および導入者は、禁止される行為を回避するため、その役割、ならびにシステムの設計、開発および実際の使用に対する管理の程度にしたがって、異なる責任を負う。各禁止事項については、バリューチェーンにおいて、誰が、特定の予防措置および軽減措置をとり、かつAIシステムの開発および使用がAI法の目的およびアプローチへの適合性を確保するのに最も適しているかを考慮し、その役割と責任をふさわしい方法で解釈されなければならない。

2.5. AI法の適用範囲からの除外

(21) AI法第2条は、AI法第5条に列挙される禁止事項の実際の適用の完全な理解のため、関連する適用範囲に対するいくつかの一般的な除外を定める。

2.5.1. 国家安全保障、防衛および軍事目的

(22) AI法第2条第3項によれば、AI法はEU法の適用範囲外の分野には適用されず、いかなる場合にも、その権限に関する任務遂行を加盟国から委ねられた主体の種類を問わず、国家安全保障に関する加盟国の権限に影響しない。AI法は、AIシステムが、「軍事目的、防衛目的、また

[14] AI法第3条（4）
[15] AI法第2条第1項(b)
[16] AI法第2条第1項(c)

は国家安全保障目的のみで上市され、サービスが開始され、または変更の有無にかかわらず使用される場合において、これらの活動を行う主体の種類を問わず」、その適用範囲から明示的にこれを除外する。したがって、その除外が適用されるかどうかは、AI システムの目的または使用次第である。これにはそのシステムを使用して活動を行う主体だけでなく、それらの権限に関する任務遂行を加盟国から委ねられた民間のオペレータも含む場合がある。

(23) 欧州司法裁判所(CJEU)によれば、「国家安全保障」とは、「国家の本質的な機能および社会の基本的利益を保護することに対する最も重要な利益をいい、国の基本的な憲法上、政治上、経済上、または社会上の構造に深刻な不安定をもたらし得る活動、特に、テロリスト活動など、社会、住民または国家そのものを直接的に脅かし得る活動の防止および処罰を含む。」[17] 国家安全保障は、たとえば、交通安全に関する活動[18]、または司法組織もしくは司法行政[19]を含まない。CJEU が述べるとおり、「これは、加盟国が、その根本的な安全保障上の利益を定義し、その内外の安全保障を確保するための適切な措置をとることであり、[・・・]国家安全保障の保護のために[…]採られた国家的措置は、EU 法の適用を不可とするものではなく、EU 法を遵守する加盟国の義務を免除するものではない」。[20]

(24) AI 法第 2 条第 3 項第 2 段落における除外の適用のため、AI システムは、軍事、防衛、または国家安全保障の目的でのみ、上市され、サービスが開始され、または使用されなければならない。AI 法前文 24 項は、「のみ」の概念がどのように解釈されるか、および当該目的で使用される AI システムが、どのような場合に、それにもかかわらず AI 法の適用範囲に入るか、さらに明確にする。

たとえば、軍事、防衛または国家安全保障の目的で上市され、サービスが開始されまたは使用された AI システムが、他の目的、たとえば、民生目的または人道目的、法執行または公安目的で（一時的または恒久的に）使用された場合、そのシステムは AI 法の適用範囲となる。その場合、AI システムを他の目的で使用する主体は、システムが既に AI 法を遵守していない限り、AI システムが AI 法を遵守していることを確保しなければならず、それは当該使用前に確認されなければならない。

(25) さらに、AI 法前文 24 項が明確にするとおり、除外された目的、すなわち軍事、防衛または国家安全保障の目的で、および民生目的または法執行目的など 1 つ以上の除外されていない目的で（いわゆる「デュアルユース」システム）、上市されまたはサービスが開始された AI システムは、AI 法の適用範囲に入る。

たとえば、ある企業が、法執行や国家安全保障を含むさまざまな目的で、RBI システムを提供する場合、その企業は、当該「デュアルユース」システムの提供者であり、AI 法における要件の遵守を確保しなければならない。

[17] 欧州司法裁判所 2020 年 10 月 6 日判決、La Quadrature du Net and Others, C-511/18, C-512/18 および C-520/18, EU : C:2020:791, 135 項 ; 欧州司法裁判所 2023 年 6 月 5 日判決、Commission v Poland, C-204/21, EU : C:2023:442, 318 項、C-439/19, 67 項 および C-306/21,40 項を参照。

[18] 欧州司法裁判所 2021 年 6 月 22 日判決、Latvijas Republikas Saeima、C-439/19、EU:C:2021:504、68 項。

[19] 欧州司法裁判所 2023 年 6 月 5 日判決、Commission v Poland、C-204/21、EU:C:2023:442、319 項。

[20] 欧州司法裁判所 2020 年 10 月 6 日判決、Privacy International、C-623/17、EU:C:2020:790、44 項。

(26)　ただし、AI システムが AI 法の適用範囲に入り得るとの事実は、それらの活動を行う主体の種類にかかわらず、国家安全保障、軍事および防衛目的でそのシステムを使用する国家安全保障、防衛および軍事活動を行う主体の能力に影響を与えるべきではない[21]。

たとえば、国家安全保障機関または民間事業者が、国家諜報機関から、国家安全保障目的（たとえば情報収集）でリアルタイム RBI システムを使用することを命じられた場合、当該使用は AI 法の範囲から除外される。

(27)　AI システムが、AI 法の適用範囲内となる法執行目的で、上市され、サービスが開始されまたは使用される場合、国家安全保障上の除外の明確な表明は、特に重要である。これは、AI 法第 5 条第 1 項(d)および(h)においてそれぞれ定められる、個別の犯罪予測および評価に関する禁止事項、ならびに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係する。警察およびその他の法執行機関は、刑事犯罪の防止、探知、捜査および訴追、または公共の安全に対する脅威の保護および防止を含む刑事罰の執行の任務を負う[22]。当該目的のために AI システムが利用される場合は常に、AI 法の適用範囲に入る。

(28)　ユーロポール、および Frontex など他の EU 安全保障機関の活動は、AI 法の適用範囲に入る。

2.5.2. 第三国との司法および法執行協力

(29)　AI 法第 2 条第 4 項によれば、AI 法は、第三国の公的機関または国際機関には適用されない。その場合、それらの公的機関または国際機関は、EU または 1 つ以上の加盟国と法執行および司法協力のための国際協力または協定の枠組みで、AI システムを使用する。ただし、当該第三国または国際機関が、個人の基本的権利および自由の保護に関し、適切な保護措置を規定することを条件とする。必要に応じ、この除外は、当該法執行および司法協力を支援する特定の任務を遂行するために、問題となる第三国によって委託された民間の主体の活動を対象とし得る[23]。同時に、適用除外について、これらの国際協力または協定の枠組みが、個人の基本的権利および自由の保護に関する適切な保護措置を含むものである必要があり、それは法執行機関および司法分野において使用される AI システムに監督権限のある市場監視当局によって評価される[24]。AI 法前文 22 項が明確にするとおり、EU 域内において AI の当該出力を使用する国内の受益当局、および EU の機関、組織、部署および事務所は、その使用が EU 法を遵守していることを確保する責任を負う。将来的に、これらの国際協定が変更され、または新たな協定が締結された場合、締約国は、これらの協定が AI 法の要件に適合するよう最大限の努力を払わなければならない。

---

[21] AI 法前文 24 項。
[22] AI 法第 3 条（46）
[23] AI 法前文 22 項参照
[24] AI 法前文 22 項および第 74 条第 8 項参照。

2.5.3. 研究開発

(30) AI 法第 2 条第 8 項によれば、AI 法は、「AI システムまたは AI モデルが上市されまたはサービスが開始される前の、AI システムまたは AI モデルに関する研究活動、テスト活動、または開発活動には適用されない」。この除外は、AI 法のマーケット・ベースの論理に従うものであり、AI システムがひとたび上市されまたはサービスが開始されれば、それらに適用される。

> たとえば、研究開発（R&D）段階において、AI 開発者は、新たな機能を実験し、またテストする自由を享受するが、消費者向けアプリケーションで使用される場合、操作的とみなされ、AI 法第 5 条第 1 項(a)の対象となり得る技術を含む可能性がある。AI 法は、その上市に先立ち、AI 技術を磨き、それらが安全性および倫理基準を満たすことを確保するために、初期段階の R&D が不可欠であることを認識し、当該実験を許容する。

(31) AI 法前文 25 項で明確にするとおり、AI 法は、イノベーション支援を目的とし、AI 技術の進歩ならびに学術の進歩およびイノベーションへの寄与における、学術研究の重要性を認識する。したがって、AI 法第 2 条第 6 項は、「学術的な研究開発目的のみにより、特別に開発され、かつサービスが開始された AI システムまたは AI モデルで、その出力を含む」ものに対する除外を規定する。

> たとえば、AI が引き起こすサブリミナルな刺激または欺瞞的な刺激に対する認知上および行動上の反応に対する研究は、将来のより安全かつ効果的な AI アプリケーションの情報を与え、人間と AI との協働に関する貴重な洞察を提供し得る。当該研究は、AI 法第 5 条第 1 項(a)の禁止にかかわらず、AI 法の適用範囲から除外されることにより、許容される。

(32) しかし、AI 法第 2 条第 8 項の除外は、AI システムが当該研究開発活動の成果として上市されまたはサービスが開始される場合、AI 法を遵守する義務を害しない[25]。AI 法の意味におけるリアルワールドテスト[26]も、当該除外の適用対象とならない。

> たとえば、催し物の期間中、路上で RBI システムを使用し、顔認識ソフトウェアをテストすることを希望する地方自治体は、リアルワールドの条件でシステムにより識別されるボランティアを募集する。リアルワールドテストは、AI 法第 2 条第 8 項の除外に該当しないことから、計画されたテストは、AI 法第 60 条および第 61 条に定めるとおり[27] 、システムが正規 AI サンドボックスにおいてテストされるか、サンドボックス外のリアルワールドの条件においてテストするための特別な体制に従ってテストされない限り、AI 法における RBI システムの要件を完全に遵守しなければならない。

---

[25] AI 法前文 25 項。
[26] AI 法第 3 条(57)によれば、「リアルワールドテスト」とは、信頼できる確実なデータを収集し、かつ AI システムが本規則の要件と適合していることを評価し検証することを目的として、実験室外またはその他のシミュレートされた環境外のリアルワールドにおいて、意図目的のために AI システムを一時的にテストすることをいう。AI 法は、本規則の意味において AI システムを上市し、サービスを開始することに該当しない、リアルワールドテストのための特別な体制を規定する。ただし、テストに参加する者からの自由なインフォームドコンセントを得ることを含め、第 57 条または第 60 条に定めるすべての要件が満たされていることを条件とする、など ; AI 法第 60 条参照。
[27] AI 法は、正規 AI サンドボックスおよびリアルワールドテストに関する、詳細かつ具体的な義務を含む。AI 法第 57 条以下参照。

(33)　いかなる場合にも、あらゆる研究開発活動（AI 法の適用範囲から除外される場合を含む）は、認知された学術研究の倫理的および専門的基準に従って実行され、適用される EU 法[28]（たとえば、引き続き適用されるデータ保護法）に従って実施されなければならない。

2.5.4. 事業性のない個人な活動

(34)　AI 法第 2 条第 10 項は、次のように規定する。AI 法は「事業性のない厳密に個人的な活動の範囲内において AI システムを使用する、自然人である導入者に課せられる義務には適用されない」。導入者の定義からは、当該活動に従事するユーザーも除外される（上記の 2.4 参照）。それを通じて自然人が定期的に経済的利益を得る活動、またはその他の職業上、業務上、貿易上、仕事上またはフリーランスの活動に関連するいかなる活動も、「事業性のある」活動とみなされなければならない。「個人的」とは、事業性がないことの修飾語であり、その者は個人的でかつ事業性がないという立場の双方において活動すべきことを意味する。したがって、たとえば、犯罪行為は純粋に個人的なものとみなされないから、除外に含まれない。

> たとえば、自宅で顔認識システムを使用している個人は（たとえば、自宅へのアクセスを管理するため、および自宅への入口の安全性を監視するため）、AI 法第 2 条第 10 項の適用除外に該当し、したがって、映像（の一部）を法執行機関に転送することが求められるとしても、AI 法に基づく導入者の義務の対象ではない。
>
> これに対し、フリーランサー、ジャーナリスト、医師など事業活動のために 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導入者の義務を遵守する必要がある。自然人によるいかなる使用も、それが事業的な立場で行動する導入者の名またはその権限の下で行われる場合、AI 法の適用範囲内にある。
>
> さらに、犯罪行為は、何ら経済的利益が求められまたは達成されなくても、厳密に個人的な活動とはみなされない。消費者保護法またはデータ保護法、および国内の行政法違反などのようなその他の違法行為は、AI 法の除外が適用されるが、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 条第 10 項の除外は、事業性のない厳密に個人的な活動のためにシステムを使用する場合の導入者の義務に関してのみ適用される。このシステム自体は、システムを上市しまたはサービスを開始する提供者、その他の職務上の導入者、および輸入者や頒布者など、他の責任ある行為者の義務に関する AI 法の適用範囲内にある。

> たとえば、感情認識システムは、自然人が事業性のない厳密に個人的な活動に使用することを意図する場合、依然として AI 法第 6 条に分類されるハイリスク AI システムであり、AI 法を完全に遵守しなければならない。同時に、事業性のない厳密に個人的な使用のために（たとえば、自閉症の人）、それを使用する導入者は、AI 法に基づく導入者の特定の義務の対象にはならず、その使用は適用範囲外となる。

[28] AI 法前文 25 項。

2.5.5. フリーでオープンソースのライセンスに基づきリリースされた AI システム

(36) AI 法第 2 条第 12 項によれば、ハイリスク AI システムとして、または第 5 条（AI に関して禁止される行為）または第 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され、またはサービスが開始される場合を除き、AI 法は、フリーでオープンソースのライセンスに基づきリリースされた AI システムには適用されない[29]。これは、上市されまたはサービスが開始される AI システムが AI 法第 5 条に基づき禁止される行為を構成する場合、AI システムの導入者はこの除外の恩恵を受けられないことを意味する。

2.6. 禁止事項とハイリスク AI システムの要件との相互作用

(37) AI 法第 5 条により AI に関して禁止される行為は、AI 法第 6 条に従ってハイリスクに分類された AI システムとの関係、特に附属書 III に列挙されているものとの関係で、考慮されなければならない[30]。これは、ハイリスクに分類される AI システムの使用が、AI 法第 5 条における禁止事項の 1 つ以上の下ですべての要件が満たされた場合、場合により、特定の状況において禁止される行為に分類される可能性があることによる。反対に、AI 法第 5 条に列挙される禁止事項から除外されるほとんどの AI システムは、ハイリスクに分類される。

たとえば、感情認識システムが、AI 法第 5 条第 1 項(f)における禁止事項の要件を満たさない場合、AI 法第 6 条第 2 項および附属書 III（1）(c)に従い、ハイリスク AI システムに分類される。同様に、信用スコアリングに使用されるもの、または健康保険もしくは生命保険のリスク評価に使用されるものなど、一定の AI ベースのスコアリングシステムは、AI 法第 5 条第 1 項(c)に列挙される禁止事項の要件を満たさない場合、ハイリスク AI システムとみなされる[31]。他の例として、AI システムが人を評価し、医療サービスや社会保障給付など、人が不可欠な公的扶助およびサービスを受ける資格があるかどうかを判断する AI システムがあり、これらはハイリスクに分類される[32]。当該システムが容認できないソーシャルスコアリングを含み、AI 法第 5 条第 1 項(c)の要件を満たす場合、それらの上市、サービス開始および使用は、EU 域内において禁止される。

そのような場合、提供者が行うリスク評価および管理、ハイリスク AI システムのその他の要件（たとえば、データガバナンス、透明性および人間による管理）の遵守、ならびに使用説明書および人間による管理に従った適切な使用に対する導入者の義務（第 26 条）の遵守、場合により、基本的権利に対する影響の分析（第 27 条）の遵守は、上市されまたは導入されたハイリスク AI システムが合法であること、および禁止される行為を構成しないことの確保に資する。

---

[29] AI 法前文 102 項は、次のように述べる。フリーでオープンソースのライセンスに基づくソフトウェアおよびデータのリリースにより、これらは「自由に共有され得、かつ、ユーザーが、これらまたはこれらの修正版を自由に参照し、使用し、変更し、再頒布できる」。

[30] このリストにおいては、生体認証に基づく AI システムが含まれ、および雇用、教育、公的および私的なサービスへのアクセス、法執行機関など、一定の領域において特定の目的で使用される AI システムも含まれる。

[31] これは、AI 法前文 58 項および附属書 III に明記されている。

[32] AI 法前文 58 項。

(38)　最後に、AI 法第 6 条第 3 項に基づき、例外的にハイリスクとみなされない AI システムは、附属書 III のハイリスクのユースケースに該当するものであっても、依然として AI 法第 5 条の禁止事項の適用範囲に入る可能性がある。結果として、AI 法第 6 条第 3 項は、AI システムがハイリスクでないとみなされる結果にのみつながる；それは、当該 AI システムを、AI 法および禁止事項の範囲から除外するものではない。

2.7. 汎用 AI システムおよび意図目的のあるシステムに対する禁止事項の適用

(39)　禁止事項はあらゆる AI システムに適用され、それは、「意図目的」[33]があるか、または「汎用」（すなわち、多様な目的に対応する能力を有すること）であるか、直接的使用または他の AI システムへ組み込むためであるかを問わない[34]。したがって、各オペレータは、AI 法の 2 つの目的を達成するためにそのリスクおよび利益とのバランスをとりつつ、AI システムの責任あるかつ安全な提供および使用を確保するため、バリューチェーンにおけるシステムに対するその役割および管理に基づき最適となる措置を講じなければならない。

(40)　したがって、導入者は、AI 法第 5 条に基づき禁止される方法において AI システムを使用しないことを求められているが、これにはシステムの提供者により実装された安全性のガードレールを回避しないことを含む。害は、AI システムの実際の使用方法から生じることがよくあるとしても、提供者は、AI 法第 5 条により禁止される方法で動作する、または直接使用される高度な合理的可能性がある汎用 AI システムを含む AI システムを、上市しまたはサービスを開始しないことに対し責任を負う[35]。これに関し、提供者は、合理的に予見し得る範囲で保護措置を講じ、そのような有害な行動および誤用を防止し、また軽減するため、効果的で検証可能な措置を講じることも求められ、当該措置は、特定の AI システムおよびその場合の状況により、実行可能で相応なものであることが求められる。導入者との契約関係において（すなわち、AI システムの使用条件において）、提供者は、禁止される行為についてその AI システムの使用を排除すること、ならびに導入者のための使用説明書および必要な人間による管理に関する適切な情報を提供することも求められる。

たとえば、チャットボットとして使用される汎用 AI システムは、重大な害を引き起こし得る、操作的および欺瞞的な技術を導入する可能性がある。AI 法第 5 条第 1 項(a)に基づき、AI システムに関して禁止される行為、ならびに操作し、欺瞞し、および重大な害を引き起こす可能性が合理的にあり得る使用を防止するため、提供者は、AI システムが上市される前に（AI 法第 5 条第 1 項(a)）、当該チャットボットがユーザーまたは他の人もしくは人のグループに重大な害を引き起こさないよう確保するため、適切かつ相応な措置（たとえば、適切な安全かつ倫理的な設計、技術的措置およびその他の保護措置の組み込み、使用の制限、透明

[33] AI 法第 3 条(12)において次のように定義される。提供者が、使用説明書、販促資料、販売資料、および指示書ならびに技術文書において伝達する情報に明示するような、特定の使用状況および使用条件を含む、提供者が AI システムの目的とする使用。
[34] AI 法第 3 条（66）参照。
[35] これは、特に、AI 法第 5 条に列挙されるすべての禁止事項における「上市」または「サービス開始」についての言及からくるもので、使用にのみ適用される第 5 条第 1 項(h)のリアルタイム RBI システムの禁止は除く。

性、ならびにユーザーの管理、 使用説明書における適切な情報）を講じることが求められる（3.2.3.c も参照）。

(41) 一定の場合、特に禁止事項がシステムの極めて特別な目的に関連している場合[36]、提供者は、他の予防措置および軽減措置を組み込む可能性が限られている場合があり、主に導入者への適切な指示および情報の提供、ならびに必要とされる人間による管理およびシステムの禁止された使用の制限に頼らなければならなくなる。必要に応じ、当該措置には、AI システムが供給される手段、および誤用の可能性に関し、提供者が自由に使うことができる情報により、その制限の遵守をモニタリングすることも含む。誤用を検知するための潜在的モニタリング措置は、導入者の活動の一般的なモニタリングと同視されるものではなく、EU 法に従うものでなければならない。

たとえば、感情を認識または推測できる汎用 AI システムは、医療上または安全上の理由による例外が適用される場合を除き、職場または教育機関において、導入者により使用されてはならない。ただし、提供者は、システムの感情認識機能が使用される特定の状況、および AI 法第 5 条第 1 項(f)における禁止の例外が適用され得るかどうかを知る立場にない可能性がある。それにもかかわらず、当該提供者は、その利用規約において当該禁止される使用を明示的に除外することができ、また導入者の手引きとなる使用説明書において適切な情報を加えることができる。提供者が、特定の導入者により特定の禁止された目的のためにシステムが誤用されていることを知った場合、たとえば、当該誤用が報告された場合、または提供者が他の方法で知った場合、適切な措置を講じることも求められるが、システムが提供者の管理下にあるプラットフォームを通じて直接操作され、提供者がチェックを実施する場合がそのような場合に該当し得る。

2.8. 禁止事項と他の EU 法との相互作用

(42) AI 法は、他の EU 法、特に基本的権利の保護、消費者保護、雇用、労働者の保護、および製品の安全に関する EU 法を害することなく、すべての分野にわたり水平的に適用される規制[37]である。AI 法は、その予防と安全の論理（AI システムは、上市され、または一定の方法により使用され得ない）を通じ、そのような法を補完し、他の法律で禁止されない特定の有害な AI 行為を取扱うことにより、追加的な保護を提供する。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、上市およびサービス開始）および導入（すなわち使用）も焦点を当てることにより、AI 法の禁止事項が、AI バリューチェーンにおけるさまざまな箇所で、AI がかかわる有害な行為に対する措置を講じることを可能とする。

(43) 同時に、AI に関する行為が他の EU 法の適用対象となる場合も、AI 法は、適用される禁止事項に影響を与えない[38]。したがって、AI システムが AI 法により禁止されていない場合であっても、その使用は、他の一次的または二次的な EU 法に基づき、禁止されまたは違法になる可能性があることにかわりはない（たとえば、データ保護法に基づき必要とされる個人データ処理の法

[36] AI 法第 5 条第 1 項(d)ないし(h)
[37] AI 法第 2 条および前文 9 項。
[38] AI 法第 5 条第 8 項。

的根拠の欠如、EU 法等により禁止される差別などのように、特定の場合において基本的権利を尊重していないことが理由となる)。したがって、AI 法における禁止事項の遵守は、依然として AI システムの提供者および導入者に適用される他の EU 法を遵守していることの十分な条件ではない。

たとえば、職場において使用される AI による感情認識システムは、医療上または安全上の理由で使用されることにより、AI 法第 5 条第 1 項(f)の禁止から除外されるが、データ保護法、職場における衛生および安全を含む雇用および労働条件に関する EU 法および国内法の対象となることにかわりはなく、これにより、当該システムの使用に関する他の制限および保護措置が想定され得る[39]。

(44) AI システムの上市または使用に関する特定の行為が他の EU 法においても対象とされる場合、AI 法は、さまざまな規定の一貫した履行を確保することを目的とする。さらに、それは、AI 法の執行責任者である所轄当局と AI 法第 77 条および AI 法のその他の規定による基本的権利を保護する機関との間の効果的な協力を可能とする。より一般的にいえば、TEU 第 4 条第 3 項に従い、関係する各種の当局は、EU 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) 禁止事項の文脈においては、AI システムが識別されまたは識別可能な自然人に関する情報(「個人データ」)を頻繁に処理するため、AI 法と EU データ保護法との相互作用が特に重要である[40]。禁止事項および状況にもよるが、当該システムに関連する最も関連性の高い法令は、個人データ処理と関連する自然人の保護およびそのデータの自由移動に関する規則(EU) 2016/679(一般データ保護規則、以下「GDPR」)、刑事犯罪の予防、捜査、探知もしくは訴追または刑事罰の執行を目的とした所轄当局による個人データ処理に関する自然人の保護、およびそのデータの自由移動に関する指令(EU)2016/680(法執行指令、以下「LED」)、および EU の機関、組織、部署および事務所のデータ保護ルールを定める規則(EU)2018/1725(以下「EUDPR」)である。AI 法第 2 条第 7 項に従い、これらの法は影響を受けることがなく、EU データ保護アキ(acquis)と一貫性がありこれを補完する AI 法と並行して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、欧州司法裁判所により明確にされ、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(たとえば、「プロファイリング」[41]の概念に関して、AI 法は同じ概念を用いるため、特に AI 法第 5 条第 1 項(d)の禁止に関連する)。

(46) 法執行目的による生体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/制限に関し、AI 法は、LED 第 10 条に対する特別法として適用され、したがって、生体データのそのような使用および関係する処理をそれのみで規制する[42]。当該文脈において、AI 法は、指令(EU)2016/680 第 8 条に基づく個人データ処理の法的根拠となることは意図されない。当該指令の他のすべての規定は、AI 法に定められた条件に加え、特に AI 法第 5 条第 1 項(h)の限定的な

[39] AI 法前文 9 項も参照。
[40] AI 法第 2 条第 7 項 ; AI 法前文 10 項も参照。
[41] また、データ保護作業部会、*規則 2016/679 のための自動化された個々の意思決定およびプロファイリングに関するガイドライン*第 29 条も参照。 WP251rev.01、2018 年 2 月 6 日、および欧州データ保護会議(EDPB)により承認。
[42] AI 法前文 38 項。

例外を条件として許容される場合、法執行目的におけるリアルタイム(RBI)システムの使用に関し適用される。より一般的には、所轄の法執行当局（すなわち、LED 第 3 条第 7 項に基づく所轄当局）が、法執行目的によりデータを処理する場合、個人データ処理についても、LED が遵守されなければならない。

(47) AI 法第 2 条第 9 項に従い、EU の消費者保護および安全に関する法律は、これらの法律の適用範囲の対象となる AI システムにも完全に適用されることにかわりはない。

たとえば、 - トレーダー(B to C の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の実行は、ケースバイケースの評価の対象となり、「不公正」と見なされることもあり得るため、消費者法（指令 2005/29/EC）に違反する可能性がある。 - AI システムが医療診断または治療目的で使用される場合、感情を推測するための AI システムの使用は、規則(EU)2017/745（医療機器規則）の遵守も求められる可能性がある。

(48) さらに、AI 法は、規則（EU）2022/2065（以下「デジタルサービス法」）によって規制されるサービスに AI システムまたはモデルを組み込む仲介サービスプロバイダに関係する義務と併せて適用される。具体的にいえば、AI 法第 2 条第 5 項が示すとおり、AI 法は、デジタルサービス法第 II 章に定める当該プロバイダの責任に関する規定の適用に影響しない。

(49) 加えて、AI 法の禁止事項は、引き起こされた害について、適用される EU または国内の不法行為法[43]に従って、提供者または導入者が負う可能性のある責任を害しない。[43]

(50) 最後に、AI 法第 5 条の禁止事項およびそれらの禁止事項に対する明示の例外は、他の EU 法に基づく義務を回避し、義務違反を正当化するために用いることはできない。

(51) AI 法は、二次的な EU 法として、EU 条約および憲章によって保証される基本的権利および自由、ならびに EU が締約国である国際条約によって保護される基本的権利および自由に照らし、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追加的な明確化は、以下の関連する項目で定められる。

2.9. AI 法第 5 条の施行

[43] （損害、責任者、過失または立証責任などに関する）責任の条件は、欧州議会および欧州理事会の製造物責任に関する 2024 年 10 月 23 日の指令(EU)2024/2853（EEA 関連文書）、OJ L、2024/2853、2024 年 11 月 18 日、または適用される国内不法行為責任法。（人工知能に対する契約に基づかない民事責任ルールの採択に関する欧州議会および欧州理事会指令のプロポーザルも参照）(AI Liability Directive) COM/2022/496 final)。

[44] EU がいまだ人権および基本的自由の保護に関する欧州条約の締約国でない場合でも、憲章第 59 条第 3 項が定めるとおり、憲章が人権および基本的自由の保護に関する欧州条約によって保障された権利に対応する権利を含む限り、これらの権利の意味および適用範囲は、当該条約に定めるそれらのものと同一である。この規定は、より広範な保護を定める EU 法を妨げるものではない。

2.9.1. 市場監視当局

(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督機関（EU の機関、事務所および組織の市場監視当局として）は、禁止事項を含む、AI システムに関する AI 法のルールの執行責任者である。このような執行は、他の EU 製造物責任法とともに、規則(EU)2019/1020 により確立された製品の市場監視およびコンプライアンスのシステム内で実行される[45]。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 に定められる。これらの当局は、自らの主導または苦情により、禁止事項に関連して執行の措置を採ることができ、その苦情は、影響を受けるすべての人、またはそのような違反を認める根拠を有する他の自然人または法人が申し立てる権限を有する[46]。加盟国は、2025 年 8 月 2 日までに、その所轄の市場監視当局を指定しなければならない。

(54) リスクを示す AI システムを取扱うための AI 法における国内レベルでの手続きは、禁止事項を執行する状況において特に重要である[47]。市場監視当局の領域を超えた国境を越える影響がある場合、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、AI システムが禁止される行為を構成するかどうかを決定する、欧州委員会によって決定された EU セーフガード手続きに従わなければならない[48] 。この手続きは、AI システムの提供者および導入者の双方に対し法的安全を提供するため、禁止事項がすべての加盟国全般に一律に適用されるよう確保することを目的とする。AI 法の統一的な適用を確保するため、国内の市場監視当局は、これらのガイドラインに従い、AI 委員会内で協力することにより、加盟国の領土を越えない類似のケースに関する禁止事項の統一的な適用にも努めなければならない[49]。

2.9.2. 罰則

(55) AI 法は、侵害の重大性に応じ、そのそれぞれの規定の違反に対し罰則を設けるにあたり、段階的なアプローチに従う。AI 法第 5 条における禁止事項の違反は、最も重大な違反を構成するとみなされ、したがって、最高額の制裁金の対象となる。AI に関して禁止される行為を行う提供者および導入者は、最大 35,000,000 ユーロの制裁金が科される可能性があり、あるいは、違反者が企業の場合、前会計年度の全世界年間売上高 7%以下のいずれか高い方を上限とする制裁金が課される可能性がある[50]。各加盟国は、AI システムの提供者および導入者として当該加盟国に設立された公的機関および組織に対し行政罰が科せられる可能性がある場合について、その範囲で、ルールを定めなければならない。禁止事項に違反した EU の機関、組織および事務所は、1,500,000 ユーロ以下の行政罰の対象となり得る[51]。

---

[45] AI 法前文 156 項も参照。
[46] AI 法第 85 条。
[47] AI 法第 79 条
[48] AI 法第 81 条
[49] AI 法第 65 条および第 66 条。
[50] AI 法第 99 条
[51] AI 法第 100 条

(56)　同じ 1 個の禁止行為が、AI 法の 2 つ以上の規定に違反することもある（すなわち、ラベル付けをしないディープフェイクは、AI 法第 5 条第 1 項(a)に基づく欺瞞的手法を構成することもあり得る)。そのような場合においては、二重処罰禁止の原則を尊重しなければならない。いかなる場合にも、AI 法第 99 条第 7 項に定める行政罰を決定するための基準が考慮されなければならない。

(57)　AI 法第 5 条の禁止事項の違反は、他者の自由を最も侵害するものであり、最高額の制裁金を科すことになるため、その適用範囲は狭く解釈されなければならない。

3. AI 法第 5 条第 1 項(a)および(b)-有害な操作、欺瞞および悪用

(58)　AI 法第 5 条第 1 項(a)および(b)における最初の 2 つの禁止事項は、AI による操作および悪用の重大な悪影響から個人および脆弱な人々を保護することを目的とする。これらの禁止事項は、自然人または人のグループの行動に著しく有害かつ重大な影響を与えるサブリミナル技術、意図的な操作技術または欺瞞的な技術を用いる AI システム（AI 法第 5 条第 1 項(a)）、または、年齢、心身障害、もしくは特定の社会的経済的状況による脆弱性につけ込む AI システム（AI 法第 5 条第 1 項(b)）を対象とする。

3.1. 理論的根拠および目的

(59)　これらの禁止事項の根底にある理論的根拠は、個人の自律性、意思決定および自由な選択を否定しかつ損なう可能性のある操作的、欺瞞的および搾取的な AI に関する行為から、個人の自律性およびウェルビーイングを保護することである[52]。当該禁止事項は、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、これはすべての基本的権利の基礎を構成するものでもあり、本質的側面として個人の自律性を含む。特に、当該禁止事項は、個人を特定の目的達成のための単なる道具におとしめる AI システムを通じた操作や悪用を防止すること、および最も脆弱かつ有害な操作および悪用に影響を受けやすい人々の保護措置となることを目的とする。著しく有害な操作的、欺瞞的および搾取的な AI に関する行為の禁止は、安全で、透明性があり、公正な信頼できる人間中心の AI システム、ならびに人類に奉仕し、および人間の働きと 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、AI 法の広範な目的と完全に一致する。

.2. AI 法第 5 条第 1 項(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 条第 1 項(a)は、次のように規定する：

1.　AI に関する以下の行為は、禁止される：
(a)　人の認識の域を超えるサブリミナル技術を用いる AI システム、または意図的な操作をしもしくは欺瞞的な技術を用いる AI システムであって、見識のある判断を下す能力を著しく

[52] AI 法前文 29 項。

損い、それによって、その人、他の人、または人のグループに対し重大な害を引き起こしまたは合理的に引き起こし得るような、別のやり方では下さなかったであろう判断を人に下させる、人または人のグループの行動を実質的に歪曲する目的または効果を伴うものを上市し、サービスを開始し、または使用すること；

(60) AI 法第 5 条第 1 項(a)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない：

(i) 当該行為は、AI システムの「上市」、「サービス開始」または「使用」を構成すること。

(ii) 当該 AI システムは、サブリミナル技術（人の認識の域を超える）、意図的な操作技術または欺瞞的な技術を導入するものであること。

(iii) 当該 AI システムによって導入される技術は、人または人のグループの行動を実質的に歪曲する目的または効果を伴うものであること。当該歪曲は、情報に基づき判断するその能力を相当に害することになり、その結果、その人または人のグループが別のやり方では下さなかったであろう判断が下されることになる。

(iv) 歪曲された行動が、その人、他の人、または人のグループに重大な害を引き起こし、または合理的に引き起こし得ること。

(61) 禁止事項が適用されるためには、4 つの要件すべてが同時に満たされなければならない。また、導入された技術、人の行動の実質的な歪曲と、その行動から生じたまたは合理的に生じ得る重大な害との間には、相当な因果関係が必要である。

(62) 第 1 の要件、すなわち AI システムの「上市」、「サービス開始」または「使用」は、既に分析した。したがって、当該禁止事項は、AI システムの提供者および導入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においてそのような AI システムを上市し、サービスを開始し、または使用してはならない。以下のセクションでは、他の 3 点の要件に焦点を当てる。

3.2.1. サブリミナル、意図的な操作技術または欺瞞的な技術

(63) AI 法第 5 条第 1 項(a)は、3 つの代替的タイプの操作的な技術を禁止する：a) 人の認識の域を超えるサブリミナル技術：b) 意図的な操作技術：c)欺瞞的な技術。AI 法第 5 条第 1 項(a)の適用範囲に入るためには、AI システムがこれらの技術の 1 つ以上を導入するものでなければならない。

a) サブリミナル技術

(64) AI 法は「サブリミナル技術」を定義していないが、AI 法第 5 条第 1 項(a)は、サブリミナル技術が意識的認識の閾値（下または上）を超えて動作すると規定する。サブリミナル技術およびその動作の方法は内に隠されているため、そのような技術は操作に対する人の合理的な防御を迂

回し、重大な倫理的懸念を生じさせかつ個人の自律性、自主性、自由な選択を損ない、その人が意識的に認識することなく判断に影響を与え得る[53]。

(65) サブリミナル技術は、当該人が、当該影響、それがどのように機能するか、または当該人の判断もしくは価値・意見の形成に及ぼす効果に気がつかないままにすることにより、行動に影響を与え得るものでなければならない。特に、サブリミナル技術は、聴覚、視覚または触覚メディアを通じてもたらされる刺激を使用することがあるが、それらは気がつくには短すぎるかまたは微妙すぎ、かつメディア広告などの他の分野においては伝統的に知られかつ禁止されている。[54] これらの刺激は、意識的には知覚されないが、なお脳により処理され、かつ行動に影響を与え得る。

サブリミナル技術の例として、次のものが含まれる（AI 法第 5 条第 1 項(a)に列挙される他のすべての要件が満たされない限り、必ずしも禁止されない）：

- ビジュアルサブリミナルメッセージ：AI システムは、ビデオ再生中に短く点滅する画像やテキストを示しまたは組み込むことができ、これらは技術的には視覚可能であるが、意識が認識するには速すぎる点滅であり、それでも態度または行動に影響を与えることができる。

- 聴覚サブリミナルメッセージ：AI システムは、意識的認識なく、リスナーに影響を与える低い音量の音または他の音で隠蔽した、音または口頭のメッセージを導入し得る。これらの音は、技術的には聞こえる範囲内にはあるが、その希薄さまたは他の音によるマスキングのため、リスナーに意識的に感知されない。

- 触覚サブリミナル刺激：AI システムは、無意識のうちに受容され、感情の状態や行動に影響を与えることができる、微妙な身体的感覚を刺激し得る。

- サブビジュアルおよびサブオーディブル・キューイング：AI システムは、ただ微妙であったり隠蔽されたりした刺激だけでなく、通常の条件の下では人間の感覚ではそれらをまったく知覚できない方法で示される刺激を導入し得る。たとえば、速すぎて人間の目が意識的に検知できない点滅する視覚刺激（たとえば、点滅する画像）または人間の耳では感知できない音量で音を再生することである。

- 埋め込み画像：AI システムは、意識的に知覚されない他の視覚コンテンツ内に画像を隠すことができ、それでも脳によって処理され、行動に影響を与え得る。

-誤導：AI システムは、多くの場合、認知バイアスや注意力の脆弱性を利用し、他のコンテンツに気づかないよう、特定の刺激またはコンテンツに注意を引くことができる。

[53] AI 法前文 29 項。
[54] 特に、視聴覚メディアサービスの提供に関する加盟国の法、規制または行政措置により定める一定の規定の調整に関する 2010 年 3 月 10 日欧州議会および欧州理事会指令 2010/13/EU (OJ L 95、2010 年 4 月 15 日、p.1)（「AVMSD」）参照。これは、視聴覚商業通信におけるサブリミナル技術を厳しく禁止する。

- 時間的操作：AI システムは、ユーザーの協働において時間の感覚を作り変えることができ、その結果、ユーザーの行動に影響を与え、不寛容や依存を引き起こす。

(66) ビッグデータ分析、ニューロ技術、ブレイン・コンピュータ・インターフェース、バーチャルリアリティのような AI および関連技術の急速な発展は、高度なサブリミナル操作のリスク、および潜在意識において人間の行動に効果的に影響を与える能力を高める。[55] AI は、台頭するマシン・ブレインインターフェース、およびドリームハッキングや脳スパイウェアのような高度な技術にも拡張できる。

たとえば、あるゲームでは AI によるニューロ技術やマシン・ブレインインターフェースを活用することができ、ユーザーは脳の活動を探知するヘッドギアでゲーム（の一部）を操作することが可能になる。AI は、ユーザーに重大な害を引き起こし得る方法により、神経データ情報から、非常に介入的かつ機微な情報（たとえば個人の銀行情報、内密の情報など）を明らかにしまたは推測するため、秘密裏に、かつユーザーが気づかないままユーザーの脳を訓練するために使用され得る。AI 法第 5 条第 1 項(a)の禁止事項は、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、プライバシーと個人の自律性を尊重し、かつ安全でセキュアな方法で設計されたマシン・ブレインインターフェースアプリケーション全般を対象とはしない。

b）意図的な操作技術

(67) 「意図的な操作技術」は AI 法には定義されていないが、個人の自律性および自由な選択を損なう方法で個人の行動に影響を与え、変更し、または制御するように設計されまたは客観的に目的とする技術として理解される。操作技術は、通常、認知バイアス、心理的脆弱性、または個人がより影響を受けやすくする状況的要因を利用するように設計される。AI システムは、その適応性により、個人の個別的状況または脆弱性に適切に呼応でき、かつ大規模な操作の効果および影響を高めることができる。操作する能力は、当該技術の性質を決定する重要な要素であるが、提供者もしくは導入者または操作技術を導入するシステム自体もまた、害を引き起こすことを意図している必要はない[56]。

(68) あらゆる操作技術が意識的認識の閾値を超えて動作するわけではないが、多くは動作し、サブリミナル技術と重なり得る。当該技術は、のちに操作的な効果をも有するからである。AI 法前文 29 項が明確にするところによれば、第 5 条第 1 項(a)の禁止事項は、たとえ個人が影響を与える試みを認識していても、個人がその操作的な効果[57]を制御しまたは抵抗できない可能性のある技術も対象とする。その結果、個人は、意図的に操作する技術の対象とならなければ、通常は行わなかったであろう行動および判断に影響されまたはそれらを強いられ、個人の自律性または自由な選択が損なわれるまでになる。

---

[55] AI 法前文 29 項参照。

[56] この文脈において、前文 28 項ならびにガイドライン 3.2.2 および 3.2.3.参照。

[57] AI 法前文 28 項。

意図的な操作技術の例として、AI システムが背景の音声または画像を導入し、気分の変化を引き起こす感覚操作があり、たとえば、不安や精神的ストレスを増大させ、ユーザーの行動に重大な害を生じさせるほどの影響を与える。

他の例は、個人のパーソナルデータに基づき非常に説得的なメッセージを作成しおよび仕立てる AI システムまたは他の個別の脆弱性を利用する AI システムが、重大な害を生じさせるまでにその行動または選択に影響を与える、パーソナライズされた操作である。

(69) 意図的な操作技術に対する禁止事項は、人間がそのようにすることを意図することなく個人を操作する AI システムも対象とする。AI 法第 5 条第 1 項(a)は、特定の技術を導入し、または特定の操作的振る舞いを示す AI システムを禁止する。したがって、このような操作的技術を導入するのは、このような方法によりシステムを設計しまたは使用した提供者または導入者ではなく、AI システムである可能性もある。

たとえば、提供者がそれを意図しているかどうかにかかわらず、AI システムは、それが基とする訓練データに操作的技術の多数の例が含まれていることから、[58]または人間のフィードバックからの強化学習が操作的技術を通じて「ゲーム化」できることから、操作的技術を学習し得る。[59]

これに対し、システムの操作的振る舞いが単なる偶発的なものであれば、合理的に重大な害を引き起こす可能性がある場合に、提供者が適切な防止措置および軽減措置を講じている限り、システムは意図的な操作技術を導入しているとみなされない（以下の 3.2.3.c 参照）。

c）欺瞞的な技術

(70) AI 法は「欺瞞的な技術」を定義していない。AI 法前文 29 項が明確にするところによれば、これらは、人が意識的に認識していない方法で、または認識している場合でも、いまだ欺瞞から逃れることもできず、またはそれらを制御しもしくは抵抗することができない方法で、自律性、自由な意思決定および選択の自由を覆しかつ損なう技術である。AI システムにより導入される「欺瞞的な技術」は、個人を欺く目的または効果を伴う虚偽または誤解を招く情報を提示すること、およびその自律性、意思決定および自由な選択を損なう方法でその行動に影響を与えることを含むと理解されなければならない。

[58] M. Carroll et al., Characterising Manipulation from AI Systems, In Equity and Access in Algorithms,Mechanisms, and Optimization (EAAMO '23), October 30-November 1, 2023, Boston, MA, USA. 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.

(71)　この文脈において、AI 法第 5 条第 1 項(a)の禁止事項と、「ディープフェイク」および特定の AI 生成テキスト公表物に公共の利益に関するラベルを付ける[60]AI 法第 50 条第 4 項の導入者の義務、ならびに人と協働する AI システムが、人間ではなく AI と協働していることを人々に知らせるように設計されていることを確保する提供者の義務[61]との相互作用が明確にされなければならない。このような可視的な開示は緩和措置を構成するが、それは AI が生成および操作したコンテンツの探知を可能にする技術的措置を含む、提供者が提供する AI システムに組み込まれた設計上の性能を通じても可能とされ得る[62]。「ディープフェイク」およびチャットボットの可視的なラベリングは、AI が生成したコンテンツが公衆に拡散した際に発生し得る欺瞞のリスクを軽減し、かつ個人の意見や信念の形成および行動に対する有害な歪曲効果のリスクも軽減する。

(72)　これに対し、AI 法第 5 条第 1 項(a)の禁止事項の適用範囲は、はるかに限定的である。たとえば、個人を欺瞞することを目的としまたは効果を持つやり方、AI システムとの協働、または特に、可視的な形で開示されていない欺瞞的な AI 生成コンテンツとの協働にさらされていなかったならば起こらなかったであろう行動を歪めることを目的とし、または効果を持つやり方で、チャットボットまたは欺瞞的な AI 生成コンテンツが虚偽の情報や誤解を招くような情報を提示する場合が対象となり得る[63]。

(73)　意図的な操作技術と同様、欺瞞的な技術の禁止は、人間がそのように意図することなく個人を操作する AI システムも対象とする（上記 3.2.1.b 項参照)。たとえば、その提供者がそのような結果を意図しているかどうかにかかわらず、AI システムは、たとえば強化学習などによって、そのために開発されたタスクの性能が向上するという理由のみで、欺瞞的な技術を学習し得る[64]。

> AI により導入され得る欺瞞的な技術の例として、合成音声で人の友人または親族になりすまし、詐欺および重大な害を引き起こす者に装おうとする AI チャットボットがある。
>
> 他の例として、いつそれが評価対象となるかを識別することを学習し、何らかの望ましくない行動を一時的に停止し、評価期間が終了するとそのような行動を再開する AI システムがある。[65] このような欺瞞的な振る舞いは特に危険であり、システムに対する外部の人間による管理を避け、合理的に重大な害を引き起こす可能性がある場合は、禁止され得る。

[60] AI 法第 50 条第 4 項
[61] AI 法第 50 条第 1 項
[62] AI 法第 50 条第 2 項
[63] AI 法第 50 条の透明性義務は、原則として、ディープフェイクやチャットボットの操作的効果を最小限に抑えることを目的とするが、情報の通知にもかかわらず、これらの欺瞞的な技術が依然として個人に重大な影響を及ぼし、かつ個人の自律性と情報に基づいた意思決定を損なうまでにその行動を歪めたりする可能性がある、実例または文脈が存在し得る。したがって、それらは偽情報または操作的な目的で悪用されてはならず、禁止の他のすべての条件（重大な害を含む）を満たす場合、なお第 5 条第 1 項（a)の禁止の対象となる場合がある。
[64] F. Ward, F. Toni, F. Belardinelli, T. Everitt, Honesty Is the Best Policy: Defining and Mitigating AI Deception (neurips.cc); Advances in Neural Information Processing Systems 36 (NeurIPS 2023); P. Park. et al. AI deception: A survey of examples, risks, and potential solutions [2406.10162] Patterns, Volume 5, Issue 5, 100988.
[65] . Lehman, J. Clune, D. Misevic, C. Adami, L. Altenberg, J. Beaulieu, et al. The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities. Artificial life, 26(2):274–306, 2020.

これに対し、偶然、虚偽または誤解を招く情報を提示し、ハルシネーション[66]を行う生成AIシステムは、生成AIの限界および技術水準を考慮すると、AI法第5条第1項(a)にいう欺瞞的な技術を導入するとはみなされ得ない。特に、システムの提供者が、システムの制限についてユーザーに適切な通知を行い、そのような結果を最小限に抑えるためにシステムに適切な保護措置を組み込んだ場合が当該ケースとなり得る。ただし、かなり有害な結果を発生させる可能性がある場合には、システムが機微な文脈（たとえば、健康、教育、選挙など）を意図するものでなく、かつ導入もされないことを条件とする（以下の3.2.3.cの理由も参照）。

d）技術の組み合わせ

(74) AI法第5条第1項(a)は、サブリミナル技術、意図的な操作技術、もしくは欺瞞的な技術、または複合的な影響を与え得るそのような技術の組み合わせに適用される。上述のとおり、意図的な操作技術は、意識的認識の閾値を超えて動作する場合には、本質的にサブリミナルにもなり得る。

(75) さらに、意図的な操作技術および欺瞞的な技術が組み合わせで適用される場合、個人の行動に重大な影響を与え、無意識の操作および誤った信念に基づく判断を下すように仕向ける可能性がある。この組み合わせは、操作的な要素が既にその認知バイアスおよび感情的な反応を用意しているため、個人が受け取った情報に疑問を投げかけ、または批判的に評価する可能性が低くなるフィードバックループを生む可能性がある。

3.2.2. 人または人のグループの行動を実質的に歪曲する目的または効果を伴うこと

(76) AI法第5条第1項(a)の禁止事項が適用されるための第3の要件として、用いられたサブリミナル技術、意図的な操作技術、または欺瞞的な技術が、「人または人のグループの行動を実質的に歪める目的または効果」を有するものでなければならない。これは、人の自律性および自由な選択が損なわれる場合に、小さな影響ではなく、行動に重大な影響を与えることを前提とする。ただし、AI法第5条第1項(a)は、重大な歪曲を生じさせる「効果」のみを有し得る行為も対象としているため、意図は必要的な要件ではない。行為の実質的な歪曲の可能性と、AIシステムにより用いられるサブリミナル技術、意図的な操作技術、または欺瞞的な技術との間には、相当な/合理的に可能性のある因果関係が存在しなければならない。

a） 「行動を実質的に歪曲すること」の概念

(77) 人または人のグループの「行動を実質的に歪曲すること」の概念は、AI法第5条第1項(a)の核である。それは、サブリミナル技術、意図的な操作技術または欺瞞的な技術の導入に関係し、

[66] 「ハルシネーション」とは、開発者がそれを意図していないにもかかわらず捏造されたまたは事実上不正確である、不要な情報を生成AIシステムが生成する場合の、生成AIシステムにおける技術的な欠陥を説明するために使用される用語である。さらに、Ji Ziwei et al., Survey of Hallucination in Natural Language Generation | ACM Computing Surveys, 55, Issue 12, Article No.: 248, Pages 1 – 38 参照

情報に基づいた判断を行うその能力を著しく損なう方法により、人々の行動に影響を与えることを可能とし、それにより、人に別のやり方では行わなかったであろう方法により行動するよう仕向けたり、または判断を下すよう仕向けたりする可能性がある。

(78) 「相当な害」とは、情報に基づきかつ自律的な判断を行う能力が著しく低下し、その結果、個人に、別のやり方では行わなかったであろう方法により行動をするよう仕向けたり、または判断を下すよう仕向けたりすることをいう。それは、軽微なまたは無視し得る影響を超え、意見形成および信念形成に関するものを含め、意思決定および自由な選択における重大な歪曲または障害を伴う。これが示唆するところは、「実質的な歪曲」が、合法的な説得を超え、ある程度の強制、操作、または欺瞞を前提とするものであり、これは禁止事項の範囲外であるということである（以下の 3.5.1 参照）。

(79) 「情報に基づいた判断」は、利用可能な選択、各選択のリスクおよび利益、AI システムがその行動に及ぼし得る影響、ならびに必要に応じ、意思決定または人の行動にとって重要なその他の文脈上の情報を含む、関連情報の理解および知識を必要とする。

(80) 「行動の実質的な歪曲」の概念の解釈については、EU 消費者保護法、特に指令 2005/29/EC（不公正商取引慣行指令または「UCPD」）から重要な示唆が得られる。UCPD は、消費者が他の方法では行わなかった取引上の判断を消費者に行わせ得る、各種の不公正で、誤解を招く、強引な商慣行を禁止する（UCPD 第 5 条ないし第 9 条）。欧州司法裁判所および UCPD に関する委員会ガイダンスによれば[67]、消費者の経済行動が歪められている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判断に影響を与える「可能性がある」（つまり、可能である）ことを立証すれば十分である。[68] 欧州司法裁判所は、また、たとえ正確な情報であっても、消費者の意思決定プロセスを歪曲する方法で提示された場合、誤解を招く可能性があることを強調する。[69] 各国の執行機関は、（具体的に）各ケースの具体的事実および状況を調査し、（抽象的に）平均的な消費者の意思決定プロセスにおける慣行の潜在的な影響を評価することを任務とする。[70] その目的のため、各国の執行機関は、欧州司法裁判所が発展させたベンチマークであり、現在 UCPD に取り入れられた「平均的な」消費者の視点を採用しなければならない[71]。

---

[67] 域内市場における不公正な B to C の商慣行に関する欧州議会および欧州理事会指令 2005/29/EC の解釈および適用に関する委員会ガイダンス（OJ C 526、2021 年 12 月 29 日、.1 頁）も参照。
[68]欧州司法裁判所 2016 年 10 月 26 日（第 5 法廷）判決。Canal Digital Danmark A/S. EU:C:2016:800 , Case C-611/14, 73 項.
[69] 司法裁判所 2013 年 12 月 19 日判決、Trento Sviluppo and Centrale Adriatica,C-281/12、EU:C:2013:859
[70] 欧州委員会通知 – 域内市場における不公正な B to C の商慣行に関する欧州議会および欧州理事会指令 2005/29/EC の解釈および適用に関するガイダンス（OJ C 526、2021 年 12 月 29 日、1 頁）。
[71] UCPD 前文 18 項および 19 項参照。「平均的な消費者」とは、社会的、文化的および言語的要因を考慮し、適度に十分な情報を有し、適度に観察力および慎重さを有する人をいう。平均的な消費者テストは、統計的なテストではない（すなわち、一定の割合の消費者がビジネス慣行によって著しく歪められたり、著しく損なわれたりしたことを証明する必要はない）。当該テストは、比例原則に基づく。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な取引の促進との間の適切なバランスをとるために、この概念を採用した。裁判所および当局は、特定のケースにおいて平均的な消費者の典型的な反応を定めるために、それらが独自に判断できる権限を行使しなければならない。UCPD ガイダンスにおいて、欧州委員会は、それらに行動観察およびその他のデータを利用するよう助言した。「Compass Banca」事件 C-646/22 が明確にするところによれば、平均的な消費者の定義は、個人の意思決定能力が認知バイアスなどの制約により損なわれる可能性を排除するものではない。欧州司法裁判所 2024 年 11 月 14 日（第 5 法廷）判決 Compass Banca SpA v Autorità Garante della Concorrenza e del Mercato (AGCM), Case C-646/22, EU:C:2024:957.

(81) AI 法第 5 条第 1 項(a)の禁止事項の文脈において、システムが重大な害を引き起こす合理的な可能性がある方法において人のグループに影響を与える場合、市場監視当局は、AI システムが用いるサブリミナル技術、意図的な操作技術、または欺瞞的な技術が、標的となるグループにおける「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つつ、各ケースの具体的事実および状況も調査しなければならない。AI 法が UCPD を補完することを意図し[72]、かつ一貫した方法により適用されなければならないことから、このような解釈は正当化されると考えられる。同時に、AI 法第 5 条第 1 項(a)は、「自然人の行動」を歪曲する可能性にも言及していること、および「平均的な」個人の視点が一定の文脈において評価するのが困難かまたは効果的でないと証明される場合（たとえば、よく仕立てあげられた、または 「個別化された」操作または特定の脆弱なグループへの有害な影響）を考慮すると、具体的なケースは、サブリミナル技術、意図的な操作技術、または欺瞞的な技術を導入する AI システムが、どの範囲で具体的なケースにおいて個人の自律性を損なうことができるかを評価することにより、およびどの範囲で重大な害を引き起こしたかまたは引き起こす可能性があるかを評価することにより、具体的な個人の視点から検証され得る。

b）第 1 のシナリオ: 行動を「実質的に歪曲する」目的を伴うことにより禁止される AI システム

(82) AI 法第 5 条第 1 項(a)は、上述の技術を導入する AI システムで、かつ第 1 のシナリオとして「人または人のグループの行動を実質的に歪曲する目的」を有する AI システムに適用される。このような目的は、AI システムの提供者または導入者により、またはシステム自体が追求し得る暗黙の目的の範囲内でシステム自体により追求され得る[73]。この目的は、AI システムの「意図目的」とは区別されなければならない（AI 法第 3 条（12））。提供者が意図した場合でも、操作的な目的は、ほとんどの場合、当該システムが提供される使用目的ではなく、さらに、提供者が提供する情報（使用説明書、販促資料または販売資料、報告書ならびに技術文書において）において透明性もなく、そのように具体化されていないことがよくある。

> たとえば、さまざまな状況において使用できるチャットボットは、短い視覚的な指示を点滅させたり、聞き取りができない聴覚信号を埋め込んだりするなどのサブリミナルメッセージ技術を使用し、または広告におけるユーザーの感情的な依存または特定の脆弱性を利用するように設計される。これらの技術は、客観的に見て、消費者の意識的な認識なく、消費者の購入判断に影響を与えること、人々に著しく有害な財務上の決定を下すことを後押しすることを目的とする設計上の特徴があることにより、ユーザーの行動を実質的に歪曲することを「目的として」導入される。
>
> 他人になりすますために AI システムを導入することは、人が効果的に騙された場合に、その人の行動を騙しかつ実質的に歪曲することを「目的として」導入された AI システムと見なすこともでき、したがって、その人の身元に関する情報に基づいた判断能力に相当な影響を与える。

---

[72] AI 法前文 29 項。

[73] AI 法第 3 条(1)参照。同条項は、AI システムが、たとえそのように明示的にプログラムされていない場合でも、内在的にも操作的または欺瞞的な目的が含まれ得るその機能を実行する際に、黙示または明示の目的を追求する可能性があるという。

双方の例において、AI 法第 5 条第 1 項(a)の他の要件を満たす場合、特に重大な害に関し、それらのシステムは禁止事項の適用範囲に入る可能性があるが、これはケースバイケースの評価が必要である。

c) シナリオ 2：行動を実質的に歪曲する「効果を伴うこと」により禁止される AI システム

(83) 提供者または導入者が人または人のグループの行動を実質的に歪曲するとの意図は、AI 法第 5 条第 1 項(a)の禁止事項が適用されるための十分条件であるが、必要条件ではない。この禁止事項は、そのような意図が存在しないが、AI システムにより導入される技術の効果が、個人的な自律性および自由な選択を損なうほど、人または人のグループの行動を実質的に歪曲し得る場合にも適用される。

(84) しかし、AI システムにより導入されるサブリミナル技術、意図的な操作技術、または欺瞞的な技術とその行動における影響との間にある、相当な/合理的にあり得る因果関係は、禁止事項が適用されるためには常に必要である。消費者保護法との関係で、これらの影響は完全に具体化される必要はないが、当該場合のすべての状況の客観的な評価および既存の科学的知見と方法、ならびに実生活においてシステムが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、個人の自律性に具体的な影響を及ぼしかつ損なう可能性があるか、または可能であるとの十分な証拠がなければならない。この文脈においては、あるシステムが、情報に基づく判断を下す個人の能力を著しく損ない、かつその自由な選択を損なう行動を引き起こしうる事実は、その条件を十分満たし、害が現実化する「タイミング」に関する考慮事項（たとえば、追加的な行動の場合）に依存しない。

たとえば、AI 駆動のウェルビーイングチャットボットは、提供者が、ユーザーが健康的なライフスタイルを維持し続けることを支援し、かつ心理的および身体的な行動についてそれぞれに合ったアドバイスを提供することを目的とする。しかし、チャットボットが、不健康な習慣を身に付けるよう、または危険な活動（たとえば休息や給水なく、過度のスポーツをする）を行うよう、個人の脆弱性を利用する場合、そこでは、一定のユーザーがそうでなければ行わなかった当該アドバイスに従うことおよび重大な害（心臓発作、または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得る。その AI システムは、たとえ提供者がこの行動および人に有害な結果をもたらすことを意図していなかったとしても、AI 法第 5 条第 1 項(a)の禁止事項に該当する。

チャットボットが個人の自律性を著しく損ない、一定のユーザーの行動を著しく有害な方法により実質的に歪曲することが可能であり、かつ提供者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および軽減措置を講じていないとの事実だけでも、禁止行為の適用に十分である（害の合理的な可能性に関連する考慮事項については、3.2.3 および 3.5 の適用範囲を参照)。

3.2.3. 重大な害を引き起こす（合理的に引き起こし得る）

(85) 最後に、AI 法第 5 条第 1 項(a)の禁止事項が適用されるためには、人または人のグループの行動を歪曲することが、その人、他の人、または人のグループに重大な害を引き起こすか、または合理的に引き起こし得る必要がある。この文脈においては、明確化が必要となる重要な概念は、禁止事項の対象となる害の種類、害の重大性の閾値、および害と操作的または欺瞞的な技術と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) 害の種類

(86) AI 法は、操作的および欺瞞的な AI システムに関連するさまざまな種類の有害な影響に対応し、影響を及ぼし得る個々の人および人のグループに対し、それぞれ異なる意味を持つ[74]。AI 法第 5 条第 1 項(a)に関連する害の主な種類は、身体的、心理的、財政的、および経済的な害を含み[75] 、これは一定の場合においてより広範な社会的な害と合成され得る[76]。

(87) 身体的な害は、人の生命および健康に対する傷害または損害、ならびに身体の属性に対する物理的損害を含む。人の生命および健康に対する身体的な害は、多くの場合、即時に、深刻かつ取り返しのつかない結果をもたらす。AI 法は、その製品安全ロジックに沿って、重大な身体的な害をもたらす AI による操作および欺瞞を禁止することを目的とする。

たとえば、AI チャットボットは、テロリストのコンテンツを宣伝したり、または一定の人物または人のグループ（すなわちマイノリティ）に対する暴力を奨励したりすることにより、ユーザーの自傷行為を助長し、または自殺するよう、もしくは他の人や人のグループを害するよう奨励する。

(88) 心理的な害は、認知的および感情的な脆弱性を悪用する操作的な技術、および重大な害を及ぼし得る方法で個人の行動に影響を与える操作的な技術を導入する AI システムの文脈において、特に関係する。心理的な害は、人の精神的健康および心理的・感情的なウェルビーイングに対する悪影響を含む。このような害は、時間の経過とともに積み重なる可能性があり、すぐには明らかにならないかもしれないが、長期にわたり深刻な結果をもたらし得ることから、特に重大である。しかし、それらを測定することはより困難であり、特に、その場合に関係するすべての状況を考慮に入れその深刻さを判断するためには、ケースバイケースの評価が必要である。

たとえば、人の会話パターン、行動および感情をまねるよう設計された AI 交流アプリケーションは、擬人化された特徴および感情的な刺激を使用し、ユーザーの感情、習性および意見に影響を与え、依存症のような行動を奨励し、それらのユーザーがサービスに感情的に依存

[74] AI 法第 5 条第 1 項(a)参照。
[75] AI 法前文 29 項。
[76] AI 法前文 28 項参照。同条項は、禁止事項がより広範な社会的な害を引き起こす可能性があり、人間の尊厳、自由、平等、民主主義、および法の支配、ならびに憲章に記されている基本的権利の尊重という EU の価値観に反する可能性があることを説明する。また AI 法第 1 条も参照。同条は、EU の価値観として民主主義および法の支配を保護することを目的とする。

するようにし向け、自殺行動や他人への危害のリスクなど、潜在的に重大な害を引き起こす。[77]

(89) 財政的および経済的な害は、財務損失、金融排除、経済的不安定を含む、さまざまな悪影響を包含し得る。

たとえば、重大な財政的な害を引き起こす詐欺的な製品を提案するチャットボット。

(90) AI 法第 5 条第 1 項(a)を適用する際に AI システムによって引き起こされた害を評価する場合、害が分離されることはほとんどなく、それらが組み合わさって出現し、複合的かつ多面的な悪影響につながることを強調することが重要である。害の組み合わせを理解することは、その重要性を効果的に評価するために重要であり（下記の 3.2.3.b も参照）、身体的、心理的、財政的および経済的な害が組み合わさる可能性と個人およびコミュニティに対する全体的な影響を悪化させる可能性があり、さらにより広範な悪影響をもたらす可能性すらある。

たとえば

- 身体的な害を引き起こす AI システムは、心理的トラウマ、ストレスおよび不安につながる可能性もあり、ならびに逆もまた然りとなり得る。たとえば、製品およびその他の AI によるアプリケーションで使用される AI システムの中毒性のある設計は、中毒性のある行動、不安およびうつを助長することにより、心理的な害につながり得る。心理的苦痛は、その後、不眠症ならびにその他のストレス関連の健康問題および身体的問題などの身体的な害をもたらし得る。

- AI によるハラスメントは、心理的苦痛と、不眠症、身体的健康状態の悪化または免疫力の低下など、ストレスの身体的症状との双方につながり得る。

- AI の使用により生じる心理的な害は、死を含む、身体的な害につながることもあり得る。たとえば、オンラインで使用される AI システムは、ハラスメント、ストーキング、ネット上のいじめ、性的恐喝を通じ、ジェンダーに基づく暴力を助長し得る。

- 個人の心理的な害、たとえば、実在の人物に成りすます AI による「ディープフェイク」の生成により、個人の意思決定、個人的自律性、自由な選択を惑わしかつ損なうことは、人のグループに対する重大な害（たとえば、ディープフェイクに描かれた被害者と同じ民族的または人種的出自または性別を共有するなど）と結びつけられ得る。

b) 害の重大性の閾値

(91) AI 法第 5 条第 1 項(a)の禁止事項は、サブリミナル技術、操作的技術、および欺瞞的な技術によって引き起こされた害が「重大」である場合にのみ適用される。AI 法は「重大な害」の概念

[77] Renwen Zhang, Han Li, Han Meng, Jinyuan Zhan, Hongyuan Gan, and Yi-Chieh Lee. 2024. The Dark Side of AI Companionship: A
Taxonomy of Harmful Algorithmic Behaviors in Human-AI Relationships. 1, 1 (November 2024), 28 pages.

を定義していないが、人および人のグループの身体的、心理的健康、または財政的および経済的利益に対する重大な悪影響を前提とするものと理解されなければならない[78]。「重大な害」の判断は事実に特有で、各ケースにおける個々の状況を慎重に検討し、ケースバイケースで評価する必要があるが、個別の影響は、各ケースにおいて常に具体的かつ重大でなければならない。

(92) 他のEU法において、「重大な害」の概念は、ハイレベルな保護および予防措置の目標により導かれる、微妙かつ状況次第の概念として使用されることもある。[79] 類推により、何がAI法第5条第1項(a)の意味における重大な害を構成するかを評価する場合、以下の重要な考慮事項を演繹し、かつ考慮することができる。

- **害の重大性**とは、重大な害について客観的かつ観察可能な効果を持つ、AIシステムを使用することにより生じたまたは合理的に生じる可能性のある害の程度をいう。この文脈においては、AIシステムの相互依存性、さまざまな種類の害の組み合わせ、および人または人のグループに対する悪影響を考慮することが特に重要である。

- **文脈および累積的影響**[80]： 既存の状態および複数の行為の累積的影響を含む、特定の状況は、害の重大性を評価する上で重要な役割を果たす。

- **規模および強度：**害の範囲および悪影響の強度は、害が重大であるかどうかを評価する上で重要である。害が多数の人々に影響を与えるかどうかも、その重要性を評価するために関連する。

- **影響を受ける人の脆弱性：**子ども、高齢者、または障害者などの一定のグループは、特定のAI システムからの害をより受けやすいことがある。
  一般的な人にはそれほど重要ではないと考えられ得るものでも、そのような脆弱なグループ、特に子どもにとっては重大で容認できないことがある。

- **持続性および可逆性：**長期間続くまたは不可逆的な害は、通常、重大な害の閾値を満たす。
  短期的かつ可逆的な影響は、それが頻繁に生じるものでない限り、あまり重要でないと考えられ得る。

(93) TFEU第191条第2項とともに「ハイレベルな保護」を確保するとのAI法の目的は、害の重大性を評価する場合における、保護に対する包括的なアプローチを示唆する。これは、人および人のグループの個人の自律性、意思決定および自由な選択を損なうことを意図するか、または損なう可能性のあるサブリミナル技術、意図的な操作技術または欺瞞的な技術を導入するAIシステムに関連する即時的かつ直接的な害と、体系的・間接的な悪影響の双方を考慮することを意味する。

---

[78] AI法前文29項。
[79] 欧州司法裁判所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:2013:220参照。
[80] AI法前文29項参照。

たとえば、AI システムによって引き起こされる合理的可能性がある重大な身体的な害には、怪我または死亡、個人の健康に対する十分に深刻な影響、または身体の属性の破壊を含む。性的虐待および性的搾取、極端な暴力的もしくはテロリストのコンテンツなどの犯罪行為を犯すことを個人に示唆する AI システム、または個人が犯罪、自傷行為、もしくは他人を害するよう奨励する AI システムは、そのような閾値に達するとみなされなければならない。

これに対し、軽微な身体的な害には、打撲や一時的な不快など、あまり深刻でない傷害が含まれる可能性があり、それは重大または持続的な結果をもたらすものではなく、したがって、AI 法第 5 条第 1 項(a)の意味における重大性の閾値に達しない。身体的な害が、特に子どもなどの脆弱なグループに関係しているかどうかは、害の規模と同様に、および心理的、金銭的など他の種類の害と複合するかどうかと同様に、評価されなければならない。これは、当該状況およびその評価の指針となる上記に示した基準を考慮し、ケースバイケースの評価を必要とすることになる。

当該システムがサブリミナル技術、意図的な操作技術、または欺瞞的な技術を導入する可能性がある場合であっても、重大な害の閾値に達しない可能性がある場合が多い（以下の 3.5 の例を参照）。

c） 害の合理的な可能性に関する因果関係および閾値

(94) 「合理的に～し得る」の概念は、AI 法第 5 条第 1 項(a)において、その人の自由な選択を損なうような方法でその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重大な害の可能性との間に、相当な/合理的にあり得る因果関係があるかどうかを決定するために用いられる。この概念は、害が引き起こされた場合だけでなく、AI 法の安全性の論理に沿って合理的に害が引き起こされ得る場合にも、禁止事項を適用することを可能とする。この文脈において、これが特に関係するのは、AI システムの提供者または導入者が、導入されたサブリミナル技術、意図的な操作技術、または欺瞞的な技術から合理的に生じ得る重大な害を合理的に予見することができるかどうか、およびそのような重大な害のリスクを回避しまたは軽減するための適切な予防措置および軽減措置を講じたかどうかを評価するためである。これは、客観的な方法で、かつ一般に受け入れられている基準（たとえば、技術的および科学的に）に従って、合理性を判断することを前提とし、これには、AI の実務慣行および発生し得る重大な害との間に相当な因果関係を確立するための合理性の基準を含む。AI システムおよびその機能の不透明性または透明性は、この因果関係に関する結論、したがって禁止事項の適用に影響し得る。

(95) 禁止され得る AI システムの提供または使用を避けるため、そのような操作的または欺瞞的な技術を導入する AI システムの提供者および導入者は、次のような適切な措置を講じることが奨励される。

1. **透明性と個人の自律性：**これは、情報に基づいた意思決定を確保するため、AI システムが、どのようにその能力および制限に関する明確な開示をし、どのように関連情報を運用するかに関する透明性を提供し；個人の自律性を尊重し、かつ潜在的に有害な方法で個人の自律性、意

思決定、および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該当することを避け；当該システムが欺瞞的ではなく、禁止事項の適用範囲外となる合法的な説得の範囲内において動作することを確保するよう、適切なユーザーコントロールおよび保護措置を組み込む（3.5.1 参照）。

2. **関連する適用法規の遵守：**多くの場合、関連する適用法規の遵守は、害のリスクを軽減することになり、その行為が意図的な操作または欺瞞的な行為を構成しないこと、および重大な害の可能性を予防するための軽減措置が講じられていることを示す（3.4 および 3.5.1 参照）。

3. **最先端の実務および業界標準：**安全かつ倫理的な AI システムおよび害を軽減するための措置の責任ある開発および使用に関する、専門的なデューデリジェンス実務および業界基準の遵守は、意図しない重大な害を未然に防ぎかつ軽減するために役立ち得る。

(96) これに対し、AI システムの外部要因から生じる個人の行動の害および歪曲、および提供者または導入者がリスクを予防しかつ軽減するための管理下になくかつ合理的に予見できない個人の行動の害と歪曲は、システムと協働する人の歪曲された行動と重大な害との間に相当な/合理的に可能性がある因果関係があるかどうかの評価には関係しない[81]。

たとえば、AI システムの提供者は、システム設計において、ならびに設計、事前テストおよびその他の相応な緩和措置を通じ人間との協働の設計において、潜在的に有害な操作的効果を評価しかつ軽減するよう試すことができる。しかし、AI システムの提供者は、人がうつ病になるかどうかや、未知のシステムとの協働を超える人の私生活において、他の外的要因により、人の行動が変化するかどうかを予見する立場にないといえる。

(97) すべての要件を満たすものではないとして禁止事項の適用範囲外となる他の例（たとえば合法的な説得の場合）は、以下の 3.5 に定める。

3.3. AI 法第 5 条第 1 項(b)における禁止事項の主な構成要素 – 脆弱性の有害な悪用

AI 法第 5 条第 1 項(b)は、次のように規定する： 1. AI に関する以下の行為は、禁止される： (b) 自然人または一定の人のグループの年齢、心身障害、または特定の社会的もしくは経済的状況に起因するなんらかの脆弱性につけこむ AI システムであって、その人または第三者に対し重大な害を引き起こしまたは合理的に引き起こし得るような、その人またはそのグループに属する人の行動を実質的に歪曲する目的または効果を伴うものを、上市し、サービスを開始し、または使用すること；

(98) AI 法第 5 条第 1 項(b)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない。

[81] AI 法前文 29 項参照。

(i) 当該行為は、AI システムの「上市」、「サービス開始」または「使用」を構成すること。
(ii) AI システムは、年齢、心身障害、または社会的もしくは経済的状況に起因する脆弱性につけ込むものであること。
(iii) AI システムによって可能となる悪用は、人または人のグループの行動を実質的に歪曲する目的または効果があること。
(iv) 歪曲された行動が、その人、他の人、または人のグループに重大な害を引き起こし、または合理的に引き起こし得るものであること。

(99) 禁止事項が適用されるためには、4 つの要件すべてが同時に満たされなければならず、悪用、その人の行動の実質的な歪曲、およびその行動から引き起こされた重大な害または合理的に引き起こされ得る重大な害との間に相当因果関係が必要である。

(100) 第 1 の要件、すなわち AI システムの「上市」、「サービス開始」、または「使用」は、既に 2.3 において分析した。第 3 および第 4 の要件は、AI 法第 5 条第 1 項(a)の禁止事項に関するセクション 3.2.2 および 3.2.3 で検討されている。次のセクションでは、上記に掲げる特定の追加的要件、すなわち脆弱性の悪用と特定の害に関連する要件に焦点を当てる。

3.3.1. 年齢、心身障害または特定の社会的経済的状況による脆弱性の悪用

(101) AI 法第 5 条第 1 項(b)の禁止事項の適用範囲に入るためには、AI システムが、年齢、心身障害または特定の社会的経済的状況により、一定の個人または人のグループに固有の脆弱性につけ込むものでなければならず、特にそれらの者が操作的および搾取的行為の影響を受けやすくするものでなければならない。

(102) AI 法は、「脆弱性」の概念を定義していない。この概念は、認知的、感情的、身体的、およびその他の形態の感受性を含む、幅広い領域のカテゴリーを包含すると理解し得るものであり、それらは、個人または人のグループが情報に基づく判断を行う能力に影響を与え、またはその他の方法でその行動に影響を与える可能性がある。AI 法第 5 条第 1 項(b)は「あらゆる」脆弱性というが、禁止事項の対象となる関係者は、年齢、心身障害、または社会的経済的状況によって定義される者に限られる。それらの者は、原則として AI の操作的または搾取的な行為を認識しまたは抵抗する能力がより制限され、保護の強化が必要である。[82] AI 法第 5 条第 1 項(b)の文言から、この感受性は、当該グループのいずれかに属する人であることに起因するものでなければならない(「due to（起因する）」)。

(103) 「悪用」とは、(そのグループの）人または他の人に対し有害な方法で客観的にそのような脆弱性を利用することと理解され、かつ禁止事項に影響されない合法的な行為から明確に区別されなければならない（3.5.2 の適用範囲外となるもの参照)。明確に定義されたこれらのグループに属する人の脆弱性の悪用は累積的であり得（「いずれか（any)」とされるため)、その組み合わ

[82] 特に、憲章第 24 条、第 25 条および第 26 条を参照。また、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。これは、AI 開発および導入における包括性と公平性を強調する。それは、子ども、高齢者、心身障害者を含む、脆弱なグループに対し、特別な注意を払うことを求める。

せにより、害を増大させる可能性のある悪化要因を構成し得ることにもなる。年齢、心身障害、または特定の社会的経済的状況によって定義されるもの以外の脆弱なグループに属する人および人のグループの脆弱性の悪用は、AI 法第 5 条第 1 項(b)の適用範囲外である。

a）年齢

（104） 年齢は、AI 法第 5 条第 1 項(b)の禁止事項の対象となる主要な脆弱性カテゴリーであり、これには若年者と高齢者双方が含まれる。この禁止事項は、AI システムが子どもおよび高齢者が持ち得る認知に関する制限およびその他の制限の悪用を防止すること、および有害で不当な影響、操作および悪用からそれらの者を保護することを目的とする。これは、AI 法[83]の目的、および子どもの安全を確保することを目的とするその他の EU および国内の法的枠組みおよび政策と軌を一にする[84]。

（105） 子ども[85]、つまり 18 歳未満の人は、その発達段階にあるため特に操作を受けやすく、何が現実のものか、および何が AI による協働の背後にある意図かを、批判的に評価しかつ理解する能力に制限がある。子どもは、その認知上および社会感情的な未熟さゆえに、AI エージェントおよびアプリケーションに対する執着の形成に特に脆弱であり、したがって、より操作、悪用、および依存的行動が起こり得る。

たとえば、
- 子どもと協働するよう設計された AI 搭載のおもちゃは、デジタル報酬および仮想の賞賛と引き換えに、家具に登る、高い棚を探索する、鋭利な物体を扱うなど、さらにリスクのある課題を達成するよう子どもに奨励することにより、子どもがおもちゃとの協働に興味を持ち続け、子どもに重大な身体的害を引き起こす可能性のある危険な行動へと子どもを押しやる。このようなシステムは、子どもの自然な好奇心および報酬に対する欲求を濫用することにより、子どもの脆弱性を悪用している。

- ゲームは AI を使用し、子どもの個別の行動および嗜好を分析し、それに基づき依存性のある強化スケジュールおよびドーパミンのようなループを通じ、個人に対応した予測不可能な報酬を生み出し、過度な遊びおよび強迫的な使用を奨励する。当該ゲームは、その長期的な結果を理解する限られた能力、プレッシャーに対する影響力、自制心の欠如、および簡単に満足する傾向を含む、子どもに固有の脆弱性を利用し、高い依存性があるように設計される。この AI による悪用は、子どもにとって深刻かつ長期的な結果となり得、それには潜在的な依存的行動、運動不足や睡眠不足による身体的な健康問題、視力低下、集中力の問題および認知能力の低下、学業成績の低下、社会的困難を含む。それは、成人となっても及び得る長期的な潜在的結果により、子どもの発育およびウェルビーイングに著しい影響を与え得る。

[83] AI 法前文 48 項が強調する点は、子どもが憲章第 24 条および国連子どもの権利条約に記される特定の権利を有することである。さらにデジタル環境に関する UNCRC の一般意見 No.25 において詳述されるところによれば、子どもの脆弱性の考慮ならびに子どものウェルビーイングに必要な保護およびケアを提供することの双方が必要である。
[84] the new European strategy for a better internet for kids (BIK+), COM/2022/212 final 参照
[85] EU 法は、一般に、18 歳未満の者を子どもとみなし、これは国連のどもの権利条約(UNCRC)と同じである。

いずれの例においても、AI 法第 5 条第 1 項(b)の禁止事項は、子どもに深刻な害を及ぼすこのような悪用および依存症のような行為のみを対象とする。利益をもたらす可能性があり、その禁止事項のすべての条件を満たさない場合に影響を与えない AI によるおもちゃ、ゲーム、学習アプリ、その他のデジタルアプリ全般を対象とするものではない。3.5 の適用範囲外も参照。

(106) 同様に、高齢者[86]は、認知能力の低下に苦しむことがあり（たとえ認知症にならないとしても)、現代の AI 技術の複雑さに苦労することもあり、そのような場合、高齢者は詐欺または強制的な戦術に脆弱となる。

たとえば、
- AI システムは、欺瞞的なパーソナライズされたオファーや詐欺で高齢者を標的にするために使用され、その認知能力の低下を利用し、他の方法では高齢者が行わなかったであろう判断に影響を与えることを目的とするが、それは高齢者に重大な経済的な害を引き起こし得る。

- 高齢者を支援することを目的とするロボットは、高齢者の脆弱な状況を利用でき、かつその自由な選択に反し一定の活動を強制し得るが、これにより高齢者のメンタルヘルスが著しく悪化し、深刻な心理的な害を引き起こし得る。

いずれの例においても、AI 法第 5 条第 1 項(b)の禁止事項は、高齢者に深刻な害を引き起こし得る搾取的な行為のみを対象とする。利益をもたらす可能性があり、その禁止行為のすべての要件を満たさない場合に影響を与えない AI によるパーソナルアシスタント、健康アプリ、および一般的な支援ロボットを対象とするものではない。3.5 適用範囲外となるものも参照。

b）心身障害

(107) AI 法第 5 条第 1 項(b)の禁止事項が保護しようとする脆弱性の第 2 のカテゴリーは、心身障害に起因するものである。その目的は、AI システムが心身障害者の認知機能ならびにその他の制限および弱点を悪用することを防止し、有害で不当な影響、操作、悪用からそれらを保護することである。

(108) 心身障害[87]は、他の障壁との相互作用により、個人が他者と平等に社会に完全かつ効果的に参加することを妨げる、広範囲の長期的な身体的、精神的、知的および感覚的障害を含む。このような脆弱性を悪用する AI システムは、他の人に比べその障害により影響されやすくまたは利用されやすい可能性がある心身障害を持つ人にとって、特に有害となり得る。

---

[86] （訳者注ー記載なし）
[87] AI 法前文 29 項の説明によれば、「心身障害」は、製品およびサービスへのアクセシビリティ要件に関する 2019 年 4 月 17 日の欧州議会および欧州理事会指令(EU)2019/882(EEA 関連テキスト)第 3 条(1)の意味の枠内で理解される。PE/81/2018/REV/1、OJ L 151、2019 年 6 月 7 日、p.70–115。

たとえば、
- 精神障害者に対しメンタルヘルス支援および対処術の提供を目的とする治療用チャットボットは、高価な医療製品を購入するように作用し、または自分や他の人に有害な行動をとるよう仕向けるため、障害者の限定的な知的能力を悪用する可能性がある。

- AI システムは、性的虐待コンテンツで、障害のある女性や少女をオンラインで識別することができ、より効果的なグルーミング行為によりそれらを標的にし、これにより操作および虐待をより受けやすくし、かつ自分自身を守る能力を低下させ、その障害および脆弱性を悪用する。

これに対し、アクセシブルな方法で設計されていない AI アプリは、心身障害者の脆弱性を特に対象とするものではなく、単に心身障害者がアクセスできないだけであるから、心身障害者の脆弱性を悪用するものとみなされない。

c）特定の社会的経済的状況

(109) AI 法第 5 条第 1 項(b)の禁止事項が保護しようとする脆弱性の第 3 のカテゴリーは、悪用に対して関係者をより脆弱にし得る特定の社会的経済的状況による脆弱性である。「特定」とは、この文脈において、独特の個別的特性として解釈されるべきでなく、むしろ特定の脆弱な社会的または経済的グループの法的地位またはメンバーシップとして解釈されるべきである。AI 法前文 29 項は、極度の貧困の中で暮らす人および民族的または宗教的マイノリティなど、限定列挙ではないそのような状況の例が含まれる。このカテゴリーは、原則として、比較的安定しかつ長期的な特性をカバーすることを目的とするが、一時的な失業、過剰な債務または移民の地位などの過渡的状況も、特定の社会的経済的状況として含まれ得る。しかし、いかなる者でも経験し得る不満または孤独などの状況は、社会的経済的観点から特定のものではないため、対象とならない（その悪用は AI 法第 5 条第 1 項(a)の対象となり得る）。

(110) 不利な社会的経済的状況にある人は、通常、一般の人よりも脆弱で、リソースが少なく、デジタルリテラシーが低く、それにより搾取的な AI の実務慣行を見分け、または抗うことが難しくなる。AI 法第 5 条第 1 項(b)は、AI 技術が、それらの脆弱性を利用することにより、既存の財務上およびその他の社会的不平等や不公正を永続させ、または悪化させないよう確保することを目的とする。

たとえば、AI 予測アルゴリズムは、低所得者の郵便番号地域に住み悲惨な財政状況にある人を、略奪的な金融商品の広告で標的にするために使用されることがあり、それにより潜在的な絶望からそのような広告に対するその感受性を悪用し、重大な経済的な害をそれらの者に引き起こし得る。

これに対し、不慮の偏見があり、かつ偏ったトレーニングデータにより社会的に不利な立場にある人に不相応な影響を与える（間接的な差別）AI システムは、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性である場合、またはそのような差別的な影

響が保護される特性と密接に相関する他のプロキシの特性（たとえば郵便番号）をターゲットにすることによるものである場合、それらの者は直接的な差別の場合のように特別にターゲットにされていないことから、人の社会的経済的脆弱性を悪用すると当然にみなされるべきではない。同時に、そのシステムが特定の社会的経済的状況にある人または人のグループを違法に差別していることを知るAIシステム提供者または導入者は、それらの者が被る可能性がある合理的に起こり得る重大な害を認識し、かつ適切な是正措置を講じていない場合も、脆弱性を悪用するとみなされなければならない（上記3.2.3.c参照）。

(111) 特定の社会的経済的状況の文脈において、人種上の出自、民族性、国籍または宗教など、EUの平等法に基づき保護される差別の理由に関連する代用となるものの関連性を考慮することが重要である。

たとえば、社会的経済的地位と民族上の出自は重なり得るものであり、社会的経済的データを利用するAIシステムは、民族的なマイノリティまたは特定の人種上の出自の人に不均衡な影響を与え得る。これは、既存の不均衡を悪化させ、かつ体系的な差別またはこれらのグループから個人を排除することの一因にもなり得る。

ただし、AI法第5条第1項(b)は、たとえば、人がどのようなブランドやモデルの電話を所有しているか、どれだけ大きい都市に住んでいるか、どの程度およびどこに旅行するかなど、特定の社会的経済的状況における脆弱なグループと間接的な相関すらない多様な変数に基づき消費者を対象とするAIシステムには適用されない。たとえこれらの特性が個人の社会的経済的状況全般を反映し得るとしても、彼らが、その禁止事項が悪用から保護しようとしている脆弱性を有する特定の社会的経済的状況にある個人であることは決定的ではない。

(112) たとえば、移民または難民など、独特の社会的文脈にある他の者は、安定した法的地位や社会的経済的安定性を欠いていることが多く、特にAIシステムにより悪用されやすいといえる。

たとえば、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するが、その中には移民も含まれ得る。チャットボットは、移民の脆弱性および不満を識別しかつ利用するが、移民は原則として脆弱かつ不安定な特定の社会的経済的状況にあり、その疑問に呼応してそれらを過激な考えに向かわせるが、それは、当該国の（一定のグループの）人に対する暴力を含む。

3.3.2. 実質的に行動を歪曲する目的または効果を伴うこと

(113) AI法第5条第1項(b)の禁止事項が適用されるための第3の要件は、上記で検討した脆弱性の悪用が、a)「人または人のグループの行動を著しく歪曲する目的」を有するか、またはb)「効果」を有するかのいずれかでなければならないことである。これは、軽微な影響または些細な影響ではなく、実質的な影響をいい、AI法第5条第1項(b)は、重大な歪曲を引き起こす「効果」を持ち得る行為のみを対象とすることから、必ずしも意図を必要とするものではない。AI法第5条第1項(a)および(b)は、同じ概念を用いていることから、同様に解釈されなければならない。したがって、3.2.2における説明は、AI法第5条第1項(b)にも関連する。唯一の顕著な違

いは、AI 法第 5 条第 1 項(a)においては、操作的な行為が「**情報に基づいた判断を行う能力を著しく損なう**」ことを必要とすることであるが、これは AI 法第 5 条第 1 項(b)には存在しない。その理由は、子どもや他の脆弱な人の特定の脆弱性は、彼らの情報に基づき判断する能力を低下させ、かつ他の大人であればそうするように自分自身を守ることができない行動をとるよう彼らに強いるからである。

3.3.3. 重大な害を引き起こす（合理的に引き起こし得る）こと

(114) 最後に、AI 法第 5 条第 1 項(b)の禁止事項が適用されるためには、脆弱な人または人のグループの行動を歪曲することが、その人または他の人に重大な害を引き起こすか、または合理的に引き起こし得るものでなければならない。AI 法第 5 条第 1 項(a)および(b)は、同じ概念を用いていることから、同様に解釈されなければならない。したがって、害の種類に関し 3.2.3 において示される説明、害の重大性の閾値、ならびに因果関係およびその合理性は、AI 法第 5 条第 1 項(b)の解釈に同じように関係する。

(115) 3.2.3 において説明したとおり、重大な害は、AI 法第 5 条第 1 項(b)の禁止事項が適用されるために合理的に生じ得ることになる身体的、心理的、財政的、および経済的な害を含む、一連の重大な悪影響を含む。子ども、高齢者、心身障害者、社会的経済的に不利な人など、脆弱なグループにとって、これらの害は、悪用に対するその感応性の高さゆえに、特に深刻かつ多面的となり得る。大人にとって許容できる害のリスクとみなされ得るものは、多くの場合、子どもおよびこれらの他の脆弱なグループにとっては容認できない害となる。したがって、予防的アプローチは、不確実性がありかつ重大な害が潜在的である場合、特に正当化される。

(116) たとえば、子どもは感受性がとても強く、説得力のあるコンテンツを批判的に評価し、または AI によるサービスに依存し続けることを目的とする一定の搾取的な行為に抵抗したりする認知上の成熟性を有しないことがある。これはまた、その価値観、信念形成の一因となることがあり、かつ潜在的に有害な方法で行動させることがある。ここでの重大な害は、身体的および心理的なもの双方であり、悪用を識別し抗うこと、および長期的な影響を与え得る発達とウェルビーイングへの有害な影響を識別し抗うことに対し、子どもに能力がないことにより悪化する。

たとえば、 - 子どもの性的虐待素材の生成（または、実在の子どもを写した既存の素材を操作し、その子どもを題材にしたさらに新しいコンテンツを作成する）に使用される AI システム。子どものグルーミングや性的強要のための策略の結果、影響を受けた子どもに深刻な害と虐待を生じさせる可能性があり、かつそれに耐えた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]。

[88] 委員会担当者作業文書：子どもの性的虐待および性的搾取ならびに子どもの性的虐待素材に対抗することに関する欧州議会および理事会の指令の提案文書に付随する影響評価報告 SWD / 2024/33 最終版。AI 生成 CSAM に関する詳細な統計を含む Internet Watch Foundation 2024 報告の統計も参照。これはここで入手できる：https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery

- AI システムは、若年のユーザーの脆弱性を標的にすることができ、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用いることは、特に若者や少女に有害である。それらは、不安症およびうつ病、身体的不満、摂食障害、ならびに場合によっては自傷行為や自殺など、メンタルヘルスの問題を含む、深刻な心理的および身体的な害を引き起こし得る。[89] これはまた、子どもの発達に長期にわたる有害な結果をもたらし得るが、これには認知上の発達と学習への障害、社会的スキルの低下、子どもの感情的および身体的なウェルビーイングに不可欠な遊戯、睡眠、対面による社会的対話のような経験の置換えを含む。[90]

- 擬人化された方法で設計され、かつ子どもとの対話において人間のような感情的な回答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助長し、関与する時間を操作し、本物の人間関係に対する子どもの理解を歪める方法で、子どもの脆弱性を悪用することができる。これは、その通常の社会的・感情的な発達および他の人間との関係、ならびに共感、感情的な調整、および社会的理解と適応性のような社会的感情的スキルを妨げ得る[91]。結果として、これは、サービスに対する子どもの不安および依存の増加、ならびに子どものウェルビーイングに対する長期的な害など、心理的な害を引き起こし得る。

上記に説明したように、複合的に重大な害を引き起こし得る AI によるサービスのような意図的で依存性がありかつ搾取的な設計における特徴は、個人の自律性および子どもの安全を尊重し、かつ AI 法第 5 条第 1 項(b)の適用範囲外にある重大な害を引き起こさないユーザーの関与を目的とする提供者および導入者の他の正当な行為とは区別されなければならない（5.3 適用範囲外となるもの参照）。

（117） 同様に、高齢者は認知機能の減退およびデジタルリテラシーの低下に直面することがあり、AI 主導の詐欺または操作的なマーケティングの主な標的にされる。この場合における害は、経済的および心理的なものであることがよくあるが、多くの高齢者が経験する欲求不満および孤立によって悪化し、それは操作的影響を増幅させるために利用され得る。

たとえば、特に、高額な医療、不必要な保険契約、または高齢者に対する欺瞞的な投資スキームを対象とする、高齢者の減退した認知上の脆弱性を悪用する AI システムは、高齢者の貯蓄への重大な損失、借金の増加、および精神的苦痛をもたらし得る。

特定の社会的経済的状況を利用し、低所得の消費者に対しより高い価格を呈示する保険などの主要なサービスにおける、AI による一定の別価格設定行為は、同じ補償範囲に対してより多くの支払いすることになるため重大な経済的負担を導き得るが、それらは影響に対してそれらを脆弱な状態のままにする。[92]

---

[89] Elizabeth J. et al, A meta-analysis of the 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 : An Experience Sampling Study among Adolescents. Media Psychology 25, 343–366 (2022).

[91] Laestadius, L., Bishop, A., Gonzalez, M., Illenčík, D. & Campos-Castillo, C. Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. New Media & Society 146144482211420 (2022) doi:10.1177/14614448221142007; Neugnot-Cerioli, M. & Laurenty, O. M. The Future of Child Development in the AI Era. Cross-Disciplinary Perspectives Between AI and Child Development Experts. 予稿は、https://doi.org/10.48550/ARXIV.2405.19275 (2024).

[92] 2023 EIOPA Consumer Trends Report、16 頁、最終段落。

(118)　心身障害者は、搾取的かつ操作的なAIシステムが重大な害を引き起こし得る脆弱なグループでもある。

たとえば、感情認識を用い精神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の利点を約する製品の購入など、精神障害者に有害な決定をさせるよう操作することもある。これは、そのメンタルヘルスの状態を悪化させ、かつ効果がなくかつ高価な製品の購入により精神障害者から経済的搾取をする可能性があり、これは精神障害者に重大な心理的および経済的な害を引き起こし得る。

(119)　社会的経済的に不利な人は、その経済的絶望および不安定な社会的状況を利用するAIシステムの影響を特に受けやすく、情報リテラシーおよびデジタルリテラシーがより低いことが多い。

たとえば、AIチャットボットは、一定の種類のコンテンツ、恐怖に基づく物語、または搾取的な申し出に対し、彼らの感受性の高まりを識別することにより、特定の社会的経済的に不利なグループを標的にして、他人への暴力や傷害の行為を行わせ得る。このシステムのターゲットを絞ったアプローチは、これらの社会的経済的に不利な人の既存の脆弱性を悪化させ、彼らの課題を深刻にする。一定の場合、これは、AI法第5条第1項(b)に基づく重大な害の閾値に達するまでに、不安、抑うつ、無力感、社会的孤立、または自傷行為および過激化の増大を引き起こし得る。

(120)　AI法第5条第1項(a)と異なり、AI法第5条第1項(b)はグループの害について明確に言及していないが、AI法前文29項は、特定の人と個人のグループの双方が被った害に対する双方の禁止事項に言及する。そこで、2つの禁止事項は、AI法の安全性の論理、ならびに年齢、心身障害、および特定の社会的経済的状況による特定の脆弱なグループに属するすべての個人を保護する第5条第1項(b)の禁止事項の目的にも沿う、一貫した方法で解釈されなければならない。したがって、外部要因であり、他の人に影響を与え得る害は、たとえ当該システムによって直接影響を受けないとしても、AI法第5条第1項(b)に基づく害の重要性の評価においても考慮されなければならない。

たとえば

- AIによる子どもの脆弱性の悪用は、メンタルヘルス問題の有病率、医療費の増加、および慢性的な健康問題による生産性の低下を含む、長期的な社会的影響を与え得る。

- 経済的に不利な人の財政的脆弱性を悪用するAIシステムは、経済的排除を招来し、かつそれら不利な人の社会的経済的困難の負のスパイラルを生じさせ得る。このような悪用は、差別および社会的不平等の永続化および悪化、ならびにこれらのグループの排除を含む、社会的構造および価値に対し広範な悪影響を及ぼす社会的な害を引き起こし得る。

- 誤情報またはヘイトスピーチにより一定の脆弱な社会的経済的グループを標的にするチャットボットは、暴力および他の人の負傷や死亡をも生じさせる社会の分裂および過激化を招来し得る。

(121) 搾取的なAIによるこれらの行為の例は、子ども、心身障害者、または特定の社会的経済的状況にある人の脆弱性を悪用することがなく、かつ重大な害を合理的に引き起こし得ない一方、適切に設計され使用された場合にそれらの人に利益をもたらすことを目的とする他の多くのAIシステムとは区別されなければならない（3.5.適用範囲外となるものも参照）。

たとえば
- 学習およびゲームにおいて子どもをサポートするAIシステム；

- パーソナルアシスタントまたは支援ロボットのような、日常生活において高齢者を支援し、かつその健康および医療を改善するAIシステム、またはデジタルスキルを向上させるAIシステム；

- 社会における社会的弱者の経済的融合およびその他の融合を支援し、それらのスキルを向上させるなどするAIシステム；

- 視聴覚障害者を支援する、または適合性のあるパーソナライズされた学習を提供するAIシステムおよびデバイス；

- 障害者が製品やサービスを使用するための障壁を取り除く、アクセシブルなソリューションを生成するAIシステム；

- 日常生活において障害者を支援し、かつその社会への融合と完全な参加を可能にするAIによる義肢など。

3.4. AI法第5条第1項(a)および(b)の禁止事項の相互作用

(122) AI法第5条第1項(a)と(b)の禁止事項との間の相互作用は、各規定が補完的な方法において適用されることを確保するため、各規定がカバーする特定の文脈の説明が必要である。

(123) AI法第5条第1項(a)の禁止事項の主な焦点は、技術の性質、特に、意識的認識の閾値の下で動作する技術、またはその他の意図的な操作技術または欺瞞的な技術に置かれる。ここで重要な要素は、影響が主に隠されているという性質、および情報に基づき自律的な判断を行うための認知上の自律性を損なうシステムにより影響を受ける個人に対する打撃である。

(124) これに対し、AI法第5条第1項(b)の禁止事項の主な焦点は、年齢、心身障害、または特定の社会的経済的状況のために特に脆弱な人の保護であり、これらの人は、原則として、内在的要因または状況的要因によりAIの悪用の影響を受けやすく、したがって、悪用に対する追加的保護が必要である。ここで重要な要素は、影響を受ける脆弱な人の特性と、それらの特定の脆弱性がAIシステムにより悪用されている事実である。

たとえば、AIシステムが高速画像フラッシュを使用し、購買の判断に影響を与える場合、操作のサブリミナルな性質により、AI法第5条第1項(a)に該当し得る。逆に、認知能力の減退

を利用し高齢者を保険加入の標的とするAIシステムは、AI法第5条第1項(b)に該当し得る。

(125) 両規定が適用されると考えられるシナリオにおける、区別の主な基準は、悪用の主たる側面である。悪用が関係者の特定の脆弱性に関係なく当てはまる場合、その操作的または欺瞞的な技術が脆弱な人の行動に及ぼす特定の影響、およびその人が経験し得る特定の害を考慮に入れ、AI法第5条第1項(a)を優先すべきである。これに代わり、AIによる操作と悪用が、その年齢、心身障害、または特定の社会的経済的状況により特定の脆弱な人のグループを標的とする場合、またはその脆弱性を悪用することを目的とする場合、AI法第5条第1項(b)が適用されなければならない。他のグループの脆弱性の悪用は、それらの人の特定の脆弱性および弱点に影響を与える場合、AI法第5条第1項(a)の一環として対象となり得る。

3.5. 適用範囲外となるもの

(126) AI法第5条第1項(a)および(b)の禁止事項が適用されるためには、上記において検討したとおり、関連規定に記載されるすべての要件が満たされなければならない。これらの要件を満たさない他のすべてのAIシステムは、以下に述べるいくつかの例のとおり、これらの禁止事項の適用範囲外である。

3.5.1. 合法的な説得

(127) 操作を説得から区別することは、AI法第5条第1項(a)の禁止事項の適用範囲を明らかにするために重要であり、この禁止事項は合法的な説得行為には適用されない。操作および説得の双方が個人の判断および行動に影響を与えるとしても、それは方法と倫理的影響において著しく異なる。

(128) 操作は、ほとんどの場合、自律性を損なう隠された技術を含み、個人が、生じる影響を完全に認識していれば、他の方法では下さなかったかもしれない判断を下すように仕向ける。これらの技術は、心理的な弱点または認知バイアスをよく利用する。これに対し、説得は、透明性および個人の自律性の尊重の範囲内において作用する。これは、理性および感情に訴える方法において議論または情報を提示することを意味するが、AIシステムの目的および機能を説明し、情報に基づいた意思決定を確保するために関連する正確な情報を提供し、情報を評価しかつ自由で自律的な選択を行うため個人の能力を支援する。

たとえば、透明性のあるアルゴリズムおよびユーザーの嗜好と制御に基づくパーソナライズされた推奨を行うAIシステムは、説得に該当する。これに対し、サブリミナルな糸口（たとえば知覚できない画像）を使用し、ユーザーの知識や理解なく特定の選択に対する影響を与えるシステムは、操作を構成する。

(129) これらの技術の目的および影響も異なる。操作は、多くの場合、個人の自律性とウェルビーイングを犠牲にし、操作者に利益をもたらすことを目的とする。これに対し、説得は、両当事者の

利害および利益が一致するよう、情報を提供しかつ説得することを目的とする。倫理的説得は、情報に基づいた選択を行う個人の自律性を尊重し、かつ脆弱性の悪用を回避する。

たとえば、透明性の高い方法により運用され、かつ顧客の感情を分析し、顧客との対話を改善し、ユーザーの知識を支援する AI システムは、説得に該当し、ユーザーの利害と一致する。これに対し、ユーザーが商品を購入する可能性が高い特定の時点で、より高額な商品を提供するために、隠された方法で消費者の感情を推測するターゲティング広告に用いられる感情認識システムは、操作に該当し、消費者に不利益をもたらすことになる。

(130) 同意は、一定の場合において重要な役割を果たす。説得的な対話において、個人は影響のある試みを認識し、自由にかつ自律的にそれを選択できる。操作的な対話において、技術またはその影響に対する認識の欠如は、選択の自由および情報に基づく自律的な意思決定を失わせる。

たとえば、サブリミナル技術の導入を通じ、ユーザーが外国語をより良くかつより早く学習できるよう支援することを目的とする AI システムは、透明性のある方法において動作し、個人の自律性を尊重し、かつシステムの使用に同意するかしないかに対するユーザーの自由かつ情報に基づいた選択を尊重する場合、操作的ではない。

(131) 法規制の枠組みの遵守もまた、合法的な説得と比較し、操作を評価することについて重要な役割を果たす。したがって、透明性、公平性、および個人の権利と自律性を支える適用される法を遵守する AI に関する行為は、AI 法の下において禁止されない可能性がより高い。

たとえば、データ処理、つまり、データ主体に提供される情報における透明性義務を求める GDPR などのデータ保護法を遵守するには、欺瞞的または操作的な言い回しを避けなければならない[93]。場合によっては、ソーシャルネットワークにおけるサービス外のユーザーのデータに基づく一定のオンラインでのパーソナライズされた広告のように、個人データ処理を合法とするため、同意が要求され得る[94]。その同意は、とりわけ、自由かつ情報に基づくものでなければならない。これらの法的基準を満たす AI システムは、合法的な説得に該当する可能性がより高い。逆に、行動に影響を与えることになるこれらの要件を回避するシステムは、操作に該当する可能性がある。

(132) 特に、AI 法前文 29 項は、AI 法第 5 条第 1 項(a)および(b)の禁止事項が、一定の条件において、医療の文脈における合法行為に影響を与えないことを明確にする。

たとえば、AI によるサブリミナル技術は、使用条件として、個人またはその法定代理人の明示の同意を得ることを含む、適用される法および医療基準に従って運用される場合、精神疾患の心理的治療または身体リハビリテーションに使用され得る。

(133) さらに、AI 法前文 29 項が明確にするとおり、広告など、一般的で合法な商慣行は、「それ自体」またはその性質上、有害な操作的、欺瞞的または搾取的な AI による行為とはみなされない。

---

[93] 欧州データ保護委員会のガイドライン、 https://www.edpb.europa.eu/system/files/2023-02/edpb_03 2022_guidelines_on_deceptive_design_patterns_in_social_media_platform_interfaces_v2_en_0.pdf、18 項。
[94] 欧州司法裁判所 2023 年 7 月 4 日判決、Meta Platforms and Others, C-252/21, ECLI:EU:C:2023:537（以下、「Meta Platforms 判決」という）。

たとえば、

-ユーザーの嗜好に基づきコンテンツをパーソナライズするためにAIを使用する広告技術は、AI法第5条第1項(a)および(b)に基づき禁止されている、有害な方法で個人の自律性をむしばみまたは脆弱性を悪用するサブリミナル技術、意図的な操作技術、または欺瞞的な技術を導入しない場合、本質的に操作的ではない。GDPR、消費者保護法および規則(EU)2022/2065（以下「DSA」）に基づく関連する義務の遵守は、このようなリスクの軽減に資する。

- オンラインで子どもの性的素材を検出するためのAIモデルおよび分類装置の有効性を訓練しかつ改善する子どもの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、反対に、オンラインでの子どもの安全の向上に不可欠な、一般的で正当な行為である。

- 金融サービス、消費者保護、データ保護および差別の禁止に関するEU法に従い、入力に顧客の年齢または特定の社会的経済的状況を用いる住宅ローンまたはローンのような銀行サービスを提供するために使用されるAIシステムは、その年齢、心身障害、または特定の社会的経済的状況により脆弱と認識される人を保護しかつ支援するように設計され、これらのグループにとってより公正で持続可能な金融サービスにも貢献する、これらのグループにとって有益なものである場合、AI法第5条第1項(b)の意味における脆弱性の悪用とはみなされない。

- ドライバーの眠気や疲労を探知し、かつ安全法に従って休息をドライバーに警告するAIシステムは有益であり、AI法第5条第1項(b)の意味における脆弱性の悪用とはみなされない。

3.5.2. 重大な害を引き起こす可能性が低い操作的、欺瞞的および搾取的なAIシステム

(134) AI法第5条第1項(a)および(b)の禁止事項が適用されるために不可欠な要件は、AIによる脆弱性の操作および悪用が重大な害を引き起こすか、または合理的に引き起こし得ることである。合理的に重大な害を引き起こす可能性がないすべての操作的、欺瞞的、および搾取的なAIアプリは原則として禁止の適用範囲外にあるが、依然として適用される他のEU法を害するものではない（以下の3.6参照）。

重大な害を引き起こす可能性がないAIシステムの例として、次のものがある：

- AIコンパニオンシップシステムは、システムをより魅力的にするため、擬人化された方法でかつ情緒的なコンピューティングで設計され、ユーザーがより効果的に関与するように仕向けるが、それらに深刻な心理的、身体的またはその他の害、不健康な愛着および依存を合理的に引き起こす方法において、他の操作的または欺瞞的な行為に該当しない。

- セラピー用チャットボットは、ユーザーをより健康的なライフスタイルに導き、かつ喫煙などの悪習慣をやめるよう、サブリミナル技術を使用する。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うユーザーが、たとえ禁煙のための努力により身体的な不快感や心理的ストレスを覚えたとしても、AIによるチャットボットに重大な害を引き起こす可能性は考えられない。このような一時的な不快感は不可避であり、ユーザーの健康に対する長期的な利益が上回る。健康的な習慣を促進することを超えて、意思決定に影響を与える隠れた試みは存在しない。

- オンライン音楽プラットフォームは、うつ気質の曲に過度にさらされるのを避ける一方、ユーザーの感情を推測し、その気分に合わせた曲を自動的に推奨するため、感情認識システムを使用する。ユーザーは音楽を聴いているだけであり、ほかに害を受け、またはうつ病や不安を生じさせるものではないため、システムが合理的に重大な害を引き起こす可能性はない。

- 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、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において使用されるAIによる操作的かつ欺瞞的な技術。これらのシステムは、意図的な操作技術（たとえば認知バイアスの利用）を導入することがあるが、行動を歪曲するというユーザーの認識はなく、有益なトレーニングと意識向上の目的で、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EU法との相互作用

(135) AI法第5条第1項(a)および(b)の禁止事項は、他のEU法を害するものではなく、補完するものである。AI法第5条第1項(a)または(b)の禁止事項に該当する同様の行為はまた、他のEU法の侵害を構成する可能性があり、AI法と他の法の双方に基づく執行の対象となり得る。これが重要であるのは、これらの法におけるそれぞれの規定が、異なる利益を保護することを目的とし、かつ異なる目的、範囲および名宛人を有するからである。これにより、有害なAIの悪用や操作から人および人のグループを保護し、EU域内において安全で信頼できるAIによるサービスおよび製品を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条第1項(a)および(b)の禁止事項は、EU消費者保護法、特にAI駆動によって行われる場合を含む、誤解を招くビジネス慣行または攻撃的なビジネス慣行から消費者を保護するUCPDの目的と密接に連携する。AI法およびUCPDは双方とも、操作的で、誤解を招き、または攻撃的なAI駆動のビジネス慣行から消費者被害を積極的に防止することを目的とする。同時に、AI法第5条第1項(a)および(b)の禁止事項は、消費者だけでなく、商業の背景を超えたさまざまな文脈においてあらゆる自然人およびその行動を保護するため、適用範囲が幅広い。AI法の対象となる害はまた、経済的な害を超えて広範囲に及ぶが、AI法は消費者保護法にない重大な害の閾値を設定している。

(137)　この禁止事項は、合法的で公正かつ透明なデータ処理に関する原則を含むEUデータ保護法とも一致するものであり、データ保護法はデータ主体の個人データを保護し、最終的にその基本的権利および自律性を保護することを目的とする。より多くの（個人）データの利用可能性およびAIシステムによる当該データ処理の可能性の高まりは、AI法第5条第1項(a)および(b)の適用範囲に入るような、有害な操作的、欺瞞的、または搾取的な行為のリスクを高める。この文脈において、たとえば、サービス外のユーザーデータに基づくパーソナライズされたプロファイリングと広告[95]のための、透明性、データの最小化、公平性、および合法性に関するデータ保護ルールの遵守は、有害なパーソナライズされた操作や悪用の回避に寄与し得る。

(138)　EUの反差別法との相互作用は、AI法第5条第1項(b)の禁止事項にも関連する。これは、年齢や心身障害による脆弱性が、人が差別されない権利を有することにより保護されるものでもある一方、社会的経済的状況は人種や民族的出身のように、他のさまざまな理由と重なり合うことによる[96]。AI法における禁止事項は、他の理由に基づく禁止事項、または重大な害を伴わずEUの反差別法により既に禁止されている差別的行為に影響しない。

(139)　AI法第5条第1項(a)および(b)の禁止事項は、オンラインプラットフォームおよび検索エンジンなどのオンライン仲介サービスを規制し、かつそれらのサービスの提供における透明性および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。特に、DSA第25条第1項は、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、ユーザーの真の意図にそぐわないと考えられる行動を強要したりしないよう確保するため、ユーザーインターフェース内のダークパターンを禁止する。このようなダークパターンは、重大な害を引き起こし得る場合、AI法第5条第1項(a)の意味における操作的または欺瞞的な技術の一例を構成すると理解されなければならない。

(140)　デジタルサービス法（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、広告の透明性を確保する義務（第26条、および大規模オンラインプラットフォームまたは大規模検索エンジンについて第38条）、レコメンダシステムの使用（第27条）、未成年者の保護（DSA第28条）を定める。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、大規模オンラインプラットフォームまたは大規模検索エンジンに分類される場合、指定されたサービスのプロバイダは、そのサービスおよびアルゴリズムシステムを含む関連するシステムの設計または機能から生じるシステミックリスクを評価しかつ軽減する追加的義務を負う（DSA第34条および第35条）。リスク評価を実行する場合、大規模オンラインプラットフォームおよび大

[95] この点において特に関連性があるのは、欧州司法裁判所(2023年7月4日（大法廷)判決Case C-252/21 Meta Platforms Inc and Others v Bundeskartellamtである。欧州司法裁判所は、とりわけ、大規模なソーシャルネットワークプラットフォームによるダイレクトマーケティング目的によるサービス外ユーザーの個人データ処理が、管理者の正当な利益のために行われたとみなされ得ると判断する。これは、そのようなユーザーの利益および基本的権利を理由として、法的根拠としてユーザーから同意なしで行うことはできない。当該場合において、特に広範な処理は、ソーシャルプラットフォームがその活動に資金提供するパーソナライズされた当該広告におけるその運営者の利益に優先する（Meta Platforms判決、115項ないし118項参照)。

[96] たとえば、人種または民族的出自に関係なく人の平等な取扱いの原則を定める2000年6月29日の欧州理事会指令2000/43/EC, OJ L 180, 19.7.2000, p. 22–26；雇用および職業における処遇の平等のための一般的枠組みを設置する2000年11月27日の理事会指令2000/78/EC, OJ L 303, 2.12.2000, p. 16–22；雇用および職業の問題における男女の機会均等と処遇の平等の原則を実施する2006年7月5日の欧州議会および欧州理事会指令2006/54/EC（改訂）, OJ L 204、2006年7月26日、p23-36。商品およびサービスに対するアクセスおよび供給における男女間の処遇の平等の原則を実施する2004年12月13日の理事会指令2004/113/EC，OJ L 373、2004年12月21日、p.37-43。

規模オンライン検索エンジンのプロバイダは、そのレコメンダシステム、広告、コンテンツ監視およびその他の関連するアルゴリズムシステムが、そのようなシステミックリスクにどのように影響するかを考慮しなければならない。このようなリスク評価ではまた、システミックリスクが、とりわけ、サービスの意図的な操作および自動化された悪用により、どのように影響を受けるかを分析しなければならない（DSA 第 34 条第 2 項および DSA 前文 83 項参照）。それにもかかわらず、AI 法第 5 条第 1 項(a)または(b)の適用範囲は、仲介サービス提供者以外の他の行為者により提供または使用され得る他のさまざまなシナリオ（たとえば、チャットボット、AI によるサービスおよび製品）を対象とする。

（141） AI 法第 5 条第 1 項(a)に従う操作的な AI 技術の禁止は、有害な AI 駆動の広告[97]、およびメディア分野で著しく有害となり得るその他の AI による操作的および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、政治広告および関連サービスの提供および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関する透明性および関連するデューデリジェンス義務を含む、統一的ルールを定める規則(EU)2024/900（政治広告規則）[99]を補完する。本規則は、オンライン政治広告の文脈における特別なカテゴリーの個人データに基づくプロファイリング、および国内ルールにより定められる投票年齢を少なくとも 1 歳下回る人を対象とするターゲティングを禁止する。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および広告配信技術は、データ主体から収集された個人データに基づき、かつその明示的な同意がある場合にのみ実行できる。追加の透明性要件も適用される。すなわち、政治広告の開示、そのような技術および主なパラメータの使用の説明、ならびに AI システムの使用に関するものを含む、関連するロジックに関する追加情報などである。当該規則[100]に従った個人データ処理に基づくターゲティング政治広告は、有権者のプロファイリング、ならびに政治広告のターゲティングおよび広告提供が合法的な説得の範囲内の運用であることを確保することに役立つ。

（143） 有害な搾取的かつ欺瞞的な AI による行為という AI 法の禁止事項は、広告および消費者保護、ならびに事業者の適正行動に関する一般的な透明性ルールを定めた他の適用される EU 法を補完する（たとえば、指令 2014/65/EU MIFID、保険販売指令(EU) 2016/97[101]、消費者信用契約指令（EU) 2023/2225、通信販売指令(EU) 2002/65)、誤解を招く比較広告に関する指令 2006/114/EC、および一般的な消費者保護基準を定める消費者権利指令(EU)2011/83)。この点に関し、欧州保険・企業年金監督機構（EIOPA）は、AI システムが可能とする場合、AI 法の適

---

[97] AVMSD 第 9 条

[98]特に、子どもの保護を改善し、より効果的にヘイトスピーチに取り組むことを目的とする、視聴覚メディアサービスの提供に関する加盟国の法、規則、または行政措置に規定された一定の規定の調整に関する 2010 年 3 月 10 日の欧州議会および欧州理事会指令 2010/13/EU（指令(EU)2018/1808 により改正された視聴覚メディアサービス指令(AVMSD)）。

[99] 政治広告の透明性およびターゲティングに関する 2024 年 3 月 13 日の欧州議会および欧州理事会規則(EU)2024/900、PE/90/2023/REV/1、OJ L、2024/900、20.3.2024。

[100] 2025 年 10 月より適用開始。

[101] 保険販売に関する 2016 年 1 月 20 日の欧州議会および欧州理事会指令(EU)2016/97(改訂)、OJ L 26、2016 年 2 月 2 日、19–59 頁。たとえば、保険販売業者に対し、その顧客の最善の利益に従い、誠実、公正、専門的に行動することとする保険販売指令第 17 条第 1 項。

用範囲に入る可能性のある、差別価格に関連する一部の不公正な搾取行為に関する監督声明を既に発出している[102]。

(144) AI 法第 5 条第 1 項(a)および(b)の禁止事項はまた、AI システムと一体化した製品の安全性を確保する上で重要な役割を果たす EU 製品安全法（たとえば医療機器、玩具、機械について）を害するものではなく、補完するものである。これは、規制対象製品に対する事前の安全要件の遵守、および身体的精神的な害をもたらす安全上のリスクを引き起こさないことを確保するための積極的なモニタリングを含む。したがって、AI システムを組み込んだこれらの製品の製造者は、そのリスク評価および安全軽減措置において、関連する EU 統一安全法の論理および適用範囲にそれが適合する範囲で、これらの禁止事項を考慮しなければならない。EU 安全法はまた、AI 法の禁止事項を補完し、重大な害を引き起こさない安全リスクに介入しかつ対応することもできる。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なり、他の分野の EU 製品安全法（第 6 条に従いハイリスクに分類されず、かつ AI 法における要件の対象となる AI システムを組み込んだ製品を含む）における特定の要件の対象とならないすべての消費者製品が、特に、消費者の身体的および精神的健康リスクについて、リスクに対応し、通常または合理的に予見し得る使用条件の下で安全であることを求める。

(145) 最後に、刑法との相互作用が重要である。AI 法第 5 条第 1 項(a)および(b)の禁止事項は、詐欺、偽造、強要などの犯罪を構成し得るまたは引き起こし得る有害な行為を防止すること、またはテロリストコンテンツ、子どもの性的虐待素材、ヘイトスピーチおよび性的にきわどいディープフェイクなどの違法コンテンツの生成および頒布を防止することを目的とする。[104] 重要なことは、域内市場法として、AI 法第 5 条第 1 項(a)および(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することであり、したがって、犯罪行為を助長しかつ隠蔽し得る禁止された当該システムへのアクセスを制限することにより、早期に害を防ぐことができる。さらに、AI 法第 5 条第 1 項(a)および(b)の禁止事項はまた、EU 法または国内法に基づき刑事犯罪と認められない他の有害な行為も対象とすることがある。

4. AI 法第 5 条第 1 項(c) - ソーシャルスコアリング

(146) AI によるスコアリングが、良い行動を導き、安全性、効率性、またはサービスの質を向上させる利益をもたらし得る一方、人々を不当に扱いまたは害し、および社会的統制や監視に相当する一定の「ソーシャルスコアリング」行為がある。AI 法第 5 条第 1 項(c)の禁止事項は、特に、データが多数の無関係な社会的文脈に由来する場合、または取扱いが社会的行動の重大性に対して不均衡な場合に、個人またはグループの社会的行動または個人的特性に基づき、個人またはグループを評価または分類し、かつ有害または不利な取扱いを導く、そのような AI による容認できない「ソーシャルスコアリング」行為を対象とする。当該「ソーシャルスコアリング」の禁止

[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] 製品の安全性全般に関する、ならびに欧州議会および欧州理事会規則 (EU) No 1025/2012 および欧州議会および欧州理事会指令 (EU) 2020/1828 を改正し、欧州議会および欧州理事会指令 2001/95/EC および理事会指令 87/357/EEC を廃止する 2023 年 5 月 10 日の欧州議会および欧州理事会規則 (EU) 2023/988 (EEA 関連法)。
[104] 女性に対する暴力および家庭内暴力に対抗する 2024 年 5 月 14 日の欧州議会および欧州理事会指令 (EU) 2024/1385、PE/33/2024/REV/1、OJ L、2024/1385、2024 年 5 月 24 日。

は、公的および私的な文脈の双方において広く適用され、特定の部門または分野に限定されない[105]。

(147) 同時に、当該禁止は、合法でEU法および国内法[106]に従った、特定の目的のために人を評価する合法的な行為、特に、それらの法律が特定の評価目的のために関連するデータの種類を指定し、その結果である人の有害または不利な取扱いが正当化され均衡を確保する場合に、影響を与えることを意図するものではない（4.3.適用範囲外となるものを参照）。

4.1. 理論的根拠および目的

(148) 「ソーシャルスコアリング」行為を可能にするAIシステムは、一定の個人およびグループに対し、それらの社会からの排除を含む、差別的かつ不公平な結果をもたらし得ることに加え、EUの価値観と相容れない社会的統制および監視の実行をもたらし得る。「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、ならびに差別されない権利および平等に対する権利、データ保護、私生活および家族生活に対する権利を含むその他の基本的権利、ならびに必要に応じ、関連する社会的および経済的権利を保護することを目的とする。また、民主主義、平等（公的および私的サービスへの平等なアクセスを含む）およびEUの価値観を保護し、かつ促進することも目的とする[107]。

4.2. 「ソーシャルスコアリング」の禁止事項の主な概念および構成要素

> AI法第5条第1項(c)は、次のように規定する：
>
> AIに関する以下の行為は、禁止される：
> (c) 以下のいずれかまたは双方の状況を導くソーシャルスコアを伴う、社会的行動に基づき、または既知の、推論される、もしくは予測される人の特徴または人格に基づき、一定期間にわたる自然人または人のグループの評価または分類を目的とするAIシステムを上市し、サービスを開始し、または使用すること：
> (i) データが当初生成されまたは収集された文脈とは無関係な社会的文脈における、一定の自然人または人のグループに対する有害または不利な取扱い；
> (ii) その社会的行動またはその重大性に照らして不当なまたは不均衡な、一定の自然人または人のグループに対する有害または不利な取扱い；

(149) AI法第5条第1項(c)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない。

(i) 当該行為が、AIシステムの「上市」、「サービス開始」または「使用」を構成すること。

---

[105] ソーシャルスコアリングの禁止は、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くAIシステムを禁止することにより、犯罪を犯す可能性のある者のリスク評価および予測にのみ適用される評価/スコアリング行為に関連するより特化したAI法第5条第1項(d)の禁止とは異なる(5を参照)。
[106] AI法前文31項。
[107] AI法前文31項。

(ii) AI システムは、以下に基づき、一定期間にわたり、自然人または人のグループの評価または分類を目的としているか、または使用するものであること：
  (a) その社会的行動；または
  (b) 既知の、推論される、もしくは予測される人の特徴または人格。
(iii) AI システムの支援で作成されたソーシャルスコアは、以下のいずれかまたは双方のシナリオにおいて、人またはグループの有害なまたは不利な取扱いを導くまたは導く可能性のあるものであること：
  (a) データが当初生成されまたは収集された文脈とは無関係の社会的文脈；および／または、
  (b) その社会的行動またはその重大性に照らして不当または不均衡な扱い。

(150) AI 法第 5 条第 1 項(c)の禁止事項が適用されるためには、3 つの要件すべてが同時に満たされなければならない。最初の要件、すなわち、AI システムの上市、サービス開始または使用は、既に 2.3 において分析した。したがって、当該禁止事項は、AI システムの提供者および導入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において、そのような AI システムを上市し、サービスを開始しまたは使用してはならない。「ソーシャルスコアリング」の禁止に関する残る基準は、以下でさらに説明しかつ分析する。

4.2.1. 「ソーシャルスコアリング」：一定期間にわたる社会的行動または人の特徴もしくは人格に基づく評価または分類

a) 自然人または人のグループの評価または分類

(151) AI 法第 5 条第 1 項(c)の禁止が適用されるための第 2 の要件は、AI システムが自然人または人のグループの評価または分類を目的とするかまたは使用すること、およびその社会的行動または人　の特徴または人格に基づいてそれらにスコアを割り当てることである。システムにより生成されるスコアは、数学的な数値　(たとえば、0 から 1 へ)、ランク付け、またはラベルなどのように、さまざまな形式をとることがある。

(152) 禁止の適用範囲は広く、公共部門および民間部門の双方における評価および分類行為を含む（4.2.3 参照)。同時に、評価または分類は、自然人または自然人のグループのみに関係するものであり、したがって、原則として法人は除外される（4.3.適用範囲外となるものを参照)。

(153) 「**評価**」とは、人または人のグループに関する**評価または判断**の何らかの形の関与を示唆するが、その年齢、性別、身長などの特性に基づく人または人のグループの単なる**分類**は、必ずしも評価を導く必要性はない[108]。したがって、「分類」の適用範囲は 「評価」よりも広く、かつ、自然人または人のグループ、およびそれらの特性または行動に関する特定の評価または判断が必

[108]作業部会第 29 条、*Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7.

ずしも関係しない基準に基づく、自然人または人のグループの他の種類の分類またはカテゴリー分けも含む。

(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、それはEUデータ保護法[109]によって規制され、特別な評価形式を構成する。AI法第5条第1項(c)においては、当該概念または当該法[110]に何ら直接的な言及はないものの、AIシステムにより個人データに基づく評価が自動化された方法で行われる場合、その規定に含まれる禁止事項およびAI法の他の禁止事項にも関係し得る[111]。**プロファイリング**とは、個人（または個人のグループ）を一定のカテゴリーまたはグループに分類するため、特に、たとえば業務遂行能力；興味；またはあり得る態度について、**分析しおよび/または予測するため**、それらに関する情報を使用し、およびそれら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自然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条第1項(c)も適用される可能性がある。

> たとえば、SCHUFA I判決において、欧州司法裁判所（CJEU）は、ドイツ内で使用される信用力スコアリングシステムについて検討した[113]。その判例において、コンピュータプログラムにより生成された「スコア」は、支払約束に見合う人の能力に関する「確率値」であり、それはCJEUにより「プロファイリング」として認められた。より具体的にいえば、そのシステムは、ローンの返済など、その人の一定の特性に基づき、ある人の将来の態度の確率（「スコア」）に関する予測を確立した。スコアの確立（「スコアリング」）は、ある人を、一定の方法で行動した同じ特性を持つ他の人のグループに割り当てることにより、同様の態度が予測され得るという仮定に基づく[114]。CJEUによれば、この行為は、GDPR第4条第4項の意味における「プロファイリング」の定義に合致する[115]。この形式のプロファイリングは、AI法第5条第1項(c)の意味におけるそれらの人の特性に基づく人の評価を構成するともみなされ得るもので、AIシステムを使用して行われる、その規定の適用に関する他の要件を満たす場合に、禁止される。

b) 一定期間にわたる

(155) AI法第5条第1項(c)の禁止事項は、評価または分類が「**一定期間**」にわたるデータに基づくことを必要とする。これは、評価が、非常に特定された個別的状況からデータや態度により1回限りまたは直ちに評価されもしくは格付けされることに限定されるべきではないことを示唆す

---

[109] GDPR第4条第4項、第22条、LED第11条参照。作業部会第29条、*Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7も参照。
[110] AI法第3条(52)は、GDPR第4条(4)の定義とクロスリファレンスする「プロファイリング」の定義が含まれる。
[111] 特に、「プロファイリング」に言及する、AI法第5条第1項(d)における個人犯罪予測の禁止、一定の場合のAI法第5条第1項(f)および(g)における感情認識および生体分類の禁止。
[112] 作業部会第29条、*Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7.
[113] 欧州司法裁判所2023年12月7日判決、SCHUFA Holding (Scoring)、C-634/21、EU:C:2023:957(以下、「SCHUFA I判決」という)。たとえば、47項。
[114] 同上、14項（独自の強調）。
[115] 同上、47項。

る。同時に、禁止の適用範囲の迂回を避けるため、当該ケースのあらゆる状況を考慮し、この要件を評価することが重要である。

たとえば、移民・亡命当局は、カメラやモーションセンサーなど、一連の監視インフラを構築した難民キャンプに、部分的に自動化された監視システムを導入する。特定の個人（移民など）が逃亡を試みる危険があるかどうかを確認するために、分析されたデータが一定期間にわたり、特定の個人を評価する場合、これは「一定期間にわたる」とみなされ、他のすべての要件が満たされる場合、AI 法第 5 条第 1 項(c)の禁止事項が適用され得る。

c）社会的行動に基づき、または既知の、推論される、もしくは予測される人の特徴または人格に基づき

(156) AI 法第 5 条第 1 項(c)に基づき禁止される「評価」および「分類」行為は、i)個人または人のグループの社会的行動、またはii)彼らの既知の、推測される、もしくは予測される人の特徴または人格、あるいは双方に関連する AI によるデータ処理（多くの場合、広範囲の）に基づくものでなければならない。当該データは、個人によって直接的に提供される場合もあれば、間接的に収集される場合、すなわち、モニタリングを通じてであったり、第三者から取得したり、または他の情報からの推論により収集される場合もある。

(157) 第 1 のシナリオに関し、「**社会的行動**」は、一般的に、行為、行動、習慣、社会内の協働などを含み得る広範な用語であり、通常は複数の情報源からのデータポイントに関連する行動を含む[116]。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、社会的および私的な文脈における個人および個人のグループの行動だけでなく、ビジネスの文脈における社会的行動、たとえば、借金の支払い、一定のサービスを使用する際の行動、公的および私的な主体、政府、警察、および法（たとえば、 人が交通ルールに従うかどうか）との関係も含まれる。複数の文脈やデータポイントからの社会的行動データは、同じ主体によって一元化された方法で収集されることもあるが、ほとんどの場合、分散された方法で収集され、さまざまな情報源から組み合わせられ、これには、増加する個人の監視およびトラッキング（いわゆる 「データベーランス」）が含まれ得る。

(158) 第 2 のシナリオは、スコアリングが**人の特徴または人**に基づく場合であり、それは特定の社会的行動的側面を含む場合とそうでない場合がある。「**人の特徴**」には、たとえば、性、性的指向または性的特性、性別、性同一性、人種、民族、家族の状況、住所、収入、世帯、職業、雇用またはその他の法的地位、職場でのパフォーマンス、経済状況、金融流動性、健康、個人的な嗜好、興味、信頼性、行動、場所または移動、債務の水準、車種など[117]、個人に関連するさまざまな情報が含まれる可能性がある。「**人格**」は、原則として、人の特徴と同義のものとして解釈されるべきであるが、人格としての個人の特別なプロフィールの作成も意味する。人格は、多くの要因にも基づくことや判断を伴うこともあり、それは個人自身、他の人によって行われることもあれば、AI システムによって生成されることもあり得る。AI 法において、人格は、人格的特性および特徴といわれることがある[118]；これらの概念は一貫して解釈されなければならない。

---

[116] AI 法前文 31 項参照。
[117] そのような特性の例をいくつか挙げる AI 法前文 42 項参照。
[118] AI 法第 5 条第 1 項(d)

(159)　「**既知の、推論される、もしくは予測される**」人の特徴または人格は、区別の必要がある各種の情報および個人データである。「**既知の特徴**」は、入力としてAIシステムに提供された情報に基づくもので、それは、ほとんどの場合、検証可能な情報である。対照的に、「**推論される特徴**」は、他の情報から推論される情報に基づくもので、通常はAIシステムによって推論が行われる。「**予測される特徴**」とは、精度100%未満のパターンに基づいて推定される特徴である。「推論される」（または演繹される）データの概念は、EU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、AI法第5条第1項(c)において使用されるそれらの概念を解釈するための着想の源となり得る[119]。これらの各種のデータ使用は、スコアリング行為の精度および公平性に対し異なる影響を与え得ることから、特に、処理が不透明であるかまたはその精度の検証がより困難なデータポイントに依存する場合、考慮され得る。

4.2.2. ソーシャルスコアは、無関係な社会的文脈における有害または不利な取扱いを導く、および/または社会的行動の重大性に照らして不当または不均衡な取扱いを導くものでなければならない

a）ソーシャルスコアと取扱いとの因果関係

(160)　AI法第5条第1項(c)の禁止事項が適用されるためには、AIシステムによりまたはAIシステムの支援により作成されたソーシャルスコアが、評価される人または人のグループにとって有害または不利な取扱いを導くものでなければならない。言い換えれば、取扱いはスコアの結果でなければならず、スコアは取扱い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現化していないが、AIシステムがそのような有害な結果を生じさせることを意図しているか、または生成できる場合にも存在し得る。これは、AI法第5条第1項(c)において禁止される行為がそのようなAIシステムの「上市」も対象とすることを考慮すると、特に重要である。

(161)　AI法第5条第1項(c)は、AIシステムによって実行される評価または分類が、有害または不利な取扱いの唯一の要因であることを要求していない。したがって、それは、他の者による評価の対象となり、または他の者による評価と組み合わされ得る、AIによるスコアリング行為も対象とする。同時に、AIの出力は、ソーシャルスコアを生成する上で、重要な役割を十分に果たすものでなければならない。たとえば、公的機関が人の信頼性を評価し、かつその出力を追加的事実の人による評価と組み合わせるためにAIシステムを導入する場合、このAIによるソーシャルスコアリング行為は、AIが生成したスコアが最終的な決定において重要な役割を十分に果たす場合にのみ、禁止の適用範囲となる。ただし、以下に説明するとおり、有害または不利な取扱いに対するその他の要件が満たされていることを条件とする（4.2.2.b参照）。

[119]作業部会第29条参照、*Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679,* WP251rev.01, 6.2.2018, p. 7以下参照。

(162) スコアは、スコアを使用する当局とは異なる組織によって生成された場合であっても、有害または不利な取扱いを導き得る[120]。たとえば、公的機関は、信用性およびリスク評価を専門とする他の企業が生成した自然人の信用性評価に対するスコアを取得することがあるが、それは各種の情報源からの個人およびその行動に関する情報に基づく。

b）無関係な社会的文脈における有害または不利な取扱い、および/または不当または不均衡な取扱い

(163) AI 法第 5 条第 1 項(c)における禁止事項の最後の要件は、ソーシャルスコアの使用が次のいずれかの有害なまたは不利な取扱いをもたらす（またはもたらし得る）ことである。

i. データが当初生成され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、または
ii. 社会的行動またはその重大性に照らして不当または不均衡であること。

(164) これらの条件は代替的関係にあり、組み合わせて適用することもできる。少なくともそれらの 1 つ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。それは、多くの AI によるスコアリングおよび評価行為がそれらを満たさず、したがって禁止事項の適用範囲外にある可能性があるからである。特に、AI によるスコアリング行為が、特定の合法的な評価目的のためである場合で、かつ評価目的に関連するとみなされるデータを特定し、かつ有害または不利な取扱いが正当化され、社会的行動に相応であることを確保する、適用される EU 法および国内法に準拠している場合には、該当しない可能性がある（セクション 4.3.適用範囲外となるものを参照）。

(165) 「**不利な取扱い**」とは、スコアリングの結果として、その人または人のグループが、必ずしも特定の害または損害がないのに、他の者と比較して不利に取扱われることをいう（たとえば、詐欺が疑われる際、人々が追加的検査のために選別されるスコアリング行為の場合）。これに対し、「**有害な**」取扱いは、その人または人のグループが、取扱いから一定の害および損失を被ることが必要である。また、不利なまたは有害な取扱いは差別的であり、かつ EU の反差別法に基づき禁止されている場合、または一定の個人やグループの排除を前提とする場合もあり得るが[121]、それは禁止事項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 条第 1 項(c)は、一定の保護されるグループ（たとえば、年齢、民族上・人種上の出自、性別、宗教など）にのみ適用される EU の反差別法を超える不公正な取扱いを対象とし得る。

シナリオ 1：無関係な社会的文脈における有害または不利な取扱い

(166) AI 法第 5 条第 1 項(c)(i)に基づく第 1 のシナリオにおける、スコアに起因する有害または不利な取扱いは、データがもともと生成されまたは収集された文脈とは**無関係の社会的文脈**で行われなければならない。これは、ソーシャルスコアにより人々が不利なまたは有害な方法で取扱われ得るだけでなく、それらの社会的行動、または既知の、推論される、もしくは予測される人の

[120] この解釈は、SCHUFA I 判決における欧州司法裁判所（CJEU）の判決との一貫性があり、CJEU は、自動化された意思決定の文脈において、最終決定を下す者以外の主体によって生成された「スコア（プロファイリングを構成する評価）は、GDPR 第 22 条に基づき、自動化された決定を構成し得ると判断した。SCHUFA I 判決、42 項ないし 51 項、および 60 項ないし 62 項参照。
[121] AI 法前文 31 項。

特徴または人格に関するデータが、スコアリングが行われるものとは無関係の社会的文脈において生成されまたは収集されることを前提とする。これらの無関係な文脈から収集されまたは生成されたデータは、その後、AI システムが、評価または分類の目的で、または人もしくは人のグループの一般的な監視を導く方法で、明らかな関連性なく、人々のスコアリングのために使用されるものでなければならない。ほとんどの場合、これは人の合理的な予想に反して生じ、かつ、EU データ保護法、ならびに評価または分類に関連しかつ必要であるとみなされるデータおよびソースの種類を特定するその他の適用されうるルールに違反して生じる。この条件が満たされるかどうかは、評価の目的およびデータが収集されかつ生成された文脈を考慮し、ケースバイケースでの評価が必要となる。

AI 法第 5 条第 1 項(c)(i)に基づき禁止される無関係な社会的文脈における有害または不利な取扱いの例

- 国税局は、AI 予測ツールを使用し、その国におけるすべての納税者の税務申告について、より厳密な調査のための税務申告を選定する。AI ツールは、特定の個人を調査対象に選別するため、たとえば、年収、資産（不動産、車など）、受益者の家族に関するデータなどの関連する変数だけでなく、たとえば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を使用する。

- 社会福祉事務所は、AI システムを使用し、たとえば、一定の国籍または民族的出自の配偶者がいること、インターネット接続があること、ソーシャルプラットフォームでの行動、または職場でのパフォーマンスなど、詐欺の評価について明らかな関係性や関連性がない社会的文脈から収集されまたは推測される特性に基づく、家族手当の受給者による詐欺の可能性を推測する[122]。対照的に、公的機関は、社会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追求のため、給付金の配分に関連し、かつ合法的に収集されたデータを、詐欺のリスクを判断するために使用し得る。

- 公的労働局は、AI システムを使用し、個人が雇用について国家的な支援の恩恵に浴するべきかどうかを判断することを目的として、面接および AI ベースの評価に基づき、失業者をスコアリングする。当該スコアは、年齢や教育などの関連する人の特徴だけでなく、婚姻の状況、慢性疾患の健康データ、依存症など、評価の目的とは何ら明らかな関連性のないデータや文脈から収集されまたは推測された変数にも基づく[123]。

これらの容認できないスコアリング行為は、EU 法および国内法に従い、特定の目的のために人を評価する合法的な行為と区別され得る。特に、そのような法が、EU 法に従い、評価の目的に関連性がありかつ必要であるとみなされるデータを特定している場合である（4.3.適用範囲外となるものを参照）。

[122] 同様の国内受給制度およびソーシャルスコアリングの比較については、D. Hadwick & S. Lan, 'Lessons to be learned from the Dutch childcare allowance scandal: A comparative review of algorithmic governance by tax administrations in the Netherlands, France and Germany' (2021) World Tax Journal, Vol. 13, Issue 4.Familiales (CNAF)参照。

[123]類似のシステムが、ポーランドにおいて、「失業者のプロファイリング」システムのために使用されたが、違憲とみなされた後に、放棄された。Szymielewicz, Profiling the unemployed in Poland.: Social and Political Implications of Algorithmic Decision Making, Fundacja Panoptykon, 2015, p. 18 を参照。

シナリオ2：社会的行動に対して不均衡となる不利なまたは有害な取扱い

(167) AI スコアリングシステムが禁止される可能性がある AI 法第 5 条第 1 項(c)(ii)に基づくもう一つの代替的シナリオは、スコアに起因する取扱いが不当または社会的行動の重大性に対して不均衡な場合である。人の社会的行動の重大性と比較した、ソーシャルスコアリングに起因する影響の重大性および関係者の基本的権利への干渉の重大性は、一般比例原則を考慮し、そのような取扱いが、追求する正当な目的に対して不均衡であるかどうかにより決定されなければならない。これにはケースバイケースの評価が必要であり、ケースのすべての関連状況、ならびに社会的行動の評価および有害な扱いの比例性に関連する公平性および社会的正義のための一般的な倫理的考察および原則を考慮しなければならない。取扱いは、たとえば正当な目的がなければ、「不当」となり得る。そのような潜在的に有害なまたは不利な取扱いを規制する特定の基準および手続きを設ける EU または国内の部門ごとの法も、この評価の一部として関連し得る。

AI 法第 5 条第 1 項(c) ii) に基づき禁止される、社会的行動と比較して不当または不均衡な取扱いとなる例

- 公的機関は、AI システムを使用し、親のメンタルヘルスおよび失業などの基準だけでなく、複数の文脈から演繹される親の社会的行動に関する情報に基づき、危険な状況にある子を早期に発見するため家族をプロファイルする。その結果であるスコアに基づき、家族が調査のために選び出され「危険な状況にある」とみなされた子がその家族から取り上げられる。それは、たまたま医師の予約を逃したり、交通違反の罰金を受けたりするなど、親による軽微な違反の場合を含む。

- 地方自治体は、AI システムを使用し、さまざまな文脈において、住民の社会的行動に関連する複数のデータポイントに基づき、住民の信頼性をスコアリングする。「信頼性が低い」とみなされた住民の生成されたスコアは、ブラックリストに用いられる。すなわち、公的給付の取消し、他の重大な懲罰的措置、および管理または監視の強化に用いられる。評価において考慮された要因の中には、不十分なボランティア活動や、期限に図書館に本を返却しない、ごみ収集日以外に路上にごみを置いておく、地方税の支払いの遅れなど、ささやかな非行がある。

特に、EU 法および国内法が、有害なまたは不利な扱いを正当化しかつ社会的行動に相応であると保証する場合、これらの容認できないソーシャルスコアリング行為は、それらの法に従い適法な特定の目的のために人を評価する合法的な行為から区別され得る（4.3 適用範囲外となるもの参照)。

(168) また、AI 法第 5 条第 1 項(c)(i)および(ii)に基づき、代替的関係にあるいずれも同時に満たし得る。

AI 法第 5 条第 1 項(c) (i) 号および(ii) に基づく、不当なまたは不均衡な取扱いの例

- 税務当局は、AI システムを使用し、低所得、二重国籍、社会的行動などの基準を使用し、「故意/重大な過失」などのカテゴリーで不正が疑われる受益者を、プロファイリングしかつ割り当てることにより、児童手当の不正を検出する。リスクスコアに基づき、受益者のファイルが検査され、かつ、多くの場合、それらの育児給付が停止される。彼らは受け取った給付金を払い戻すよう通知を受け取り、もはや標準的な債権回収契約の対象とはならない。このようなスコアリングは、多くの家族に多額の負債を負わせることになり、個人およびや個人のグループに対する不当な、差別的な、有害な取扱いにつながり[124]、多くの家族を深刻な経済的困難に至らせる。

- 公的機関は、AI システムを使用し、指標中、インターネット接続、家族の状況、または受益者の教育レベルを、不正リスクの識別要因として考慮し、学生の住居付与プロセスにおける不正を抑制するが、それは関連性がなく、正当化もされない。

- 政府は、社会的な人間関係、オンライン活動、買い物の習慣、および期限どおり請求書を支払うかなど、生活のさまざまな側面における市民の行動に基づき市民を監視しおよび評価点をつける、包括的な AI ベースのシステムを導入する。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、ローンの金利がより高くなり、旅行、アパートの賃貸、さらには仕事を見つけることさえも困難になることに直面する。当該システムは、個人に対する過度の監視、およびソーシャルスコアを決定するために使用される社会的行動とは無関係の文脈における有害な取扱いにつながる一方（たとえば、雇用機会がソーシャルメディアの活動によって影響を受ける）、軽微な違反に対する過度の罰則を科すことにもつながる（たとえば、比較的軽微な犯罪に対する社会的および金銭的不利益）。

特に、EU 法および国内法が、有害なまたは不利な扱いを正当化しかつ社会的行動に相応であると保証する場合、これらの容認できないソーシャルスコアリング行為は、それらの法に従い適法な特定の目的のために人を評価する合法的な行為から区別され得る（4.3 適用範囲外となるもの参照）。

(169) AI 法第 5 条第 1 項(c)に基づく禁止事項には、一定の個人または人のグループに対し、賞や優遇措置が与えられる場合も含まれ得るが、これは他の個人に対するより不利な取扱いを前提とする（たとえば、雇用サポートプログラムの場合、住宅や強制移動のための（非）優先順位付け）。

4.2.3. 公人または私人によって提供されまたは使用されているかどうかに関係なく

(170) 既に指摘したとおり、AI 法第 5 条第 1 項(c)は、AI システムまたはスコアが公人または私人により提供されまたは使用されるかどうかにかかわらず、AI による容認できないソーシャルス

[124] オランダの育児給付スキャンダルの同様の例については、 Belastingdienst treft 232 gezinnen met onevenredig harde actie, 27.11.2019, (オランダ語)を参照。オランダの裁判所は、2020 年、「Systeem Risico Indicatie(SyRi)」が違法であると判断した。Geen powerplay maar fair play. Onevenredig harde aanpak van 232 gezinnen met kinderopvangtoeslag, 2017, p. 32 も参照。

コアリング行為を禁止する。公共部門におけるスコアリングは、力の不均衡および公共サービスへの依存関係により、人々に非常に重大な結果を及ぼす可能性があるが、同様に、企業やその他の主体によるスコアリング行為が増加している民間部門においても、有害な結果は生じ得る。

たとえば、
- 保険会社は、その生命保険の候補者の適格性の決定とは無関係で、かつそのような保険に支払う保険料の価格を決定するために使用される支出および財務情報を銀行から収集する。AI システムは、この情報を分析し、それに基づき、特定の個人または顧客のグループに対して契約を拒否するかどうか、生命保険料をより高く設定するかどうかを推奨する。

- 民間の信用調査機関は、AI システムを使用し、人々の信用力を判断し、かつ人の無関係な特徴に基づき個人が住宅ローンを組むべきかどうかを決定する。

特に、EU 法および国内法が、有害なまたは不利な扱いを正当化しかつ社会的行動に相応であると保証する場合、これらの容認できないソーシャルスコアリング行為は、それらの法に従い適法な特定の目的のために人を評価する合法的な行為から区別され得る（4.3 適用範囲外となるもの参照)。

(171) 所轄の市場監視当局がチェックする場合、それぞれその責任の範囲内において、次のことを証明するのは、AI システムの提供者および導入者である。証明の対象には、AI システムの機能の透明性を備えていること、スコアが使用される社会的文脈に関係するデータのみが評価または分類の目的で処理されていること、それらのデータが合法的に収集されたこと、システムが意図したとおりに動作していること、および結果として生じる有害または不利な取扱いが正当化され、社会的行動に相応であることを確保し、かつデータおよびデータソースの種類に関する情報を提供することを含む、AI に関する行為が合法かつ正当化されることである。適用される法の遵守、およびシステムに組み込まれ、その動作中に適用される適切で相応な保護措置は、適用の禁止を回避するのに役立つことになる一方、合法的かつ有益な目的（プロセスの有効性、サービスの質、安全性の向上など）で人の評価または分類するための AI システムの使用を可能にする(4.3. 適用範囲外となるものを参照)。

(172) ハイリスク AI システムの要件の遵守（たとえば、不可欠の公共サービスおよび給付、信用スコアリングおよび信用力評価、移住などの分野における）は、提供者や導入者がその義務（たとえば、 リスク管理、透明性、データガバナンス、基本的権利に対する影響評価、人間による管理、モニタリングなど）をそれぞれ履行する場合に考慮すべき、許容できないソーシャルスコアリング行為を、これらのハイリスク分野における評価および分類のために使用される AI システムが構成しないよう確保するために、役立ち得る。

4.3. 適用範囲外となるもの

(173) AI 法第 5 条第 1 項(c)の禁止事項は、自然人または人のグループのスコアリングにのみ適用され、したがって、いくつかの場合において、当該スコアが個人に直接的に影響を与え得るとしても（たとえば、予算の割り当ての場合における地方公共団体内のすべての市民)、評価が人の特

徴もしくは人格、または個人の社会的行動に基づいていない法人のスコアリングは、原則として、除外される。ただし、法人が、自然人のグループの社会的行動または人の特徴もしくは人格に基づき自然人のグループの評価または分類を集計した総合スコアに基づき評価され、かつ、このスコアが、それらの者に直接的に影響する場合（たとえば、会社のすべての従業員、その行動が評価される特定の学校の学生）、当該行為は、他のすべての要件が満たされている場合、AI 法第 5 条第 1 項(c)の適用範囲に入る可能性がある。これは、ケースバイケースの評価による。

(174) 「確率的価値」および予測としての AI ベースのソーシャルスコアリングは、サービスの質を評価するユーザー（オンラインカーシェアリングプラットフォームのドライバーや、宿泊施設のためのオンラインプラットフォームのホストなど）による個別の格付けからも区別されなければならない。AI 法第 5 条第 1 項(c)のすべての要件を満たす個人の評価または分類のために、データが AI システムにより他の情報と組み合わされかつ分析されない限り、このような格付けは、必ずしも AI が関与するわけではない個別の人のスコアの単なる集計である。

(175) さらに、自然人のスコアリングは常に禁止されるわけではなく、上記の分析のとおり、AI 法第 5 条第 1 項(c)のすべての要件が累積的に満たされるという限られた場合においてのみ禁止され、特に、AI 法前文 31 項は、禁止が「EU 法および国内法に従って、ある特定の目的で実施される自然人の合法的な評価に影響を与えてはならない」という。たとえば、クレジットスコアリング、リスクスコアリング、保証は、金融業および保険業のサービスの不可欠な側面である。このような行為は、その他の適法な行為（すなわち、サービスの質および効率改善のため、より効率的な支払請求の処理を確保するため、特定の従業員評価を行うため、詐欺の予防および検知のため、法執行のため、またはオンラインプラットフォーム上でのユーザーの行動をスコアリングするためなど）と同様、AI 法ならびにその他の適用される EU 法および EU 法を遵守する国内法に従って合法的に行われる場合、それ自体は禁止されない。

(176) 言い換えれば、合法的な方法によりソーシャルスコアを生成する目的で、かつ当該スコアのために使用される個人データが収集された関連する文脈において特定の目的で、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有害または不利な取扱いが正当化され、かつ社会的行動の重大性に相応である限り、禁止されない[125]。

(177) たとえば、信用スコアリング、マネーロンダリング防止などの分野におけるような、セクター別の EU 法の遵守は、評価の特定の正当な目的に関連しかつ必要なものとして使用され得るデータの種類を特定し、および、その取扱いが正当化されかつ社会的行動に相応なものであることを確保するものであり、したがって、AI に関する行為が AI 法第 5 条第 1 項(c)の禁止事項の適用範囲外とすることを確保し得る。

AI 法第 5 条際 1 項(c)の適用範囲外となる、EU 法および国内法に従った合法的なスコアリング行為の例：

[125] AI 法前文 31 項。

- 債権者または信用情報機関が顧客の財政上の信用力または未払い債務を評価するために使用する金融の信用スコアリングシステムは、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き、クレジットスコアを提供し、またはその信用力の評価を決定するが、それが、信用スコアリングの正当な目的に関連する場合であって、かつ、信用力の評価において消費者の公正な取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、AI 法第 5 条第 1 項(c)の適用範囲外である。

- 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、当該評価が、たとえばサービスの文脈における取引上の行動やメタデータ、過去の履歴、および詐欺のリスクを判断するために客観的に関連性がある情報源からの他の要因など、関連するデータに基づく場合であって、かつ詐欺的行為の結果として有害な取扱いが正当化され相応である場合には、これらの行為は禁止による影響を受けない。

- テレマティクスデバイスを通じて収集された情報で、運転手がスピード違反をしていること、または安全運転をしていないことを示すものは、保険契約者のハイリスクな運転行為に関連するテレマティクスベースの料金を提案する保険会社が使用すると、その運転行為により引き起こされる事故のリスクが高いことにより、その保険契約者の保険料を引き上げることに用いられる可能性がある。ただし、保険料の増加は、運転手の危険な行為に相応なものであることを条件とする。

- AI システムの正当な意図目的に関連しかつ必要なデータの収集および処理（たとえば、患者を診断するため、さまざまな情報源から収集される健康および統合失調症のデータ）は、特に、それが関連性のある必要なデータを処理し、概して一定の自然人に対する有害または不利となる不当な取扱いを伴わないから、AI 法第 5 条第 1 項(c)の適用範囲外である。

- 評価の文脈および目的に関連するデータに基づき、そのサービスの安全上の理由によりユーザーをプロファイリングするオンラインプラットフォームは、評価がユーザーの非行の重大性に不相応な有害な取扱いをもたらさない場合、AI 法第 5 条第 1 項(c)の適用範囲外である。

- AI によるターゲティング商業広告は、関連データ（たとえば、ユーザーの好み）に基づくもので、消費者保護、データ保護、デジタルサービスに関する EU 法に従って行われるものであり、かつユーザーの社会的行動の重大性に不相応な有害または不利な取扱（搾取的で不公平な差別価格決定など）をもたらすものではない場合、適用対象外である。

-強制移動や雇用に関する決定のために、難民キャンプにおいて収集されたデータ（たとえば、行動に関するコンプライアンス）を用いる AI システムは、このデータが評価の目的に関連するものであって、かつ適用される移民に関する EU 法に基づく手続が、当該取扱いが正

[126]特に、消費者信用契約に関する、および指令 2008/48/EC を廃止する 2023 年 10 月 18 日指令（EU）2023/2225 ならびに 2020 年 5 月 29 日以降の融資開始とモニタリングに関する欧州銀行監督局のガイドライン、EBA/GL/2020/06 参照。

当化されかつ相応であることを確保するために実施されるものである限り、禁止事項に影響しない。

- オンラインショッピングプラットフォームによるAIによるスコアリングは、多くの購入履歴があり、返品率が低いユーザーに対して、たとえば、返品申請プロセスの迅速化や返品なしの返金などの特典を提供するが、当該利点が正当化され、かつ肯定的な行動に対する報酬に相応なものであり、他のユーザーが引き続き通常の返品プロセスにアクセスできるのであれば、AI法の第5条第1項(c)の適用範囲外となる。

- 警察およびその他の法執行機関がさまざまな状況から個人の社会的行動に関するデータを収集するAIによる個人の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防止、探知、訴追および処罰の特定の目的に関連している場合であって、有害な取扱いがEUおよび各国の犯罪および警察に関する実体法および手続法に従って正当化され、かつ相応である場合、AI法第5条第1項(c)の適用範囲外となる。この文脈においては、AI法第5条第1項(d)の禁止事項を考慮することも重要であり、これはプロファイリングまたは人格的特徴の評価のみに基づくべきでない人が犯罪を犯す可能性のAIによるリスク評価および予測について、追加的かつより具体的な条件を課す（5参照）。

4.4. 他のEU法との相互作用

(178) 提供者および導入者は、その活動において使用される何らかの特定のAIスコアリングシステムに、他の適用可能なEU法および国内法が適用されるかどうか、特に、特定の評価目的に関係がありかつ必要であるとして使用され得るデータの種類を厳密に規制する、より具体的な法があるかどうか、および正当化されかつ公正な取り扱いを確保するためのより具体的なルールおよび手続きがあるかどうかについて、慎重に評価しなければならない。

(179) B to Cの関係においてトレーダーとして行動する民間の当事者のAIによるソーシャルスコアリング行為は、EU消費者保護法、すなわち不公正なB to Cの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、専門的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構成員の経済的行動を著しく歪める、または著しく歪める可能性のある商慣行を禁止する（UCPD第5条）。また、消費者の取引決定における商慣行の影響をケースバイケースで評価するならば、スコアリング行為は誤解を招き得る（UCPD第6条から第7条）。

(180) ソーシャルスコアリングは、公的または私的な当事者によるものであっても、たとえば、処理の法的根拠（合法性）、データ保護の原則（データの最小化および必要性、公平性、透明性など）、およびその他の義務（必要に応じ、自動化された個々の意思決定のみによるとのルールを含む）に関し、EUのデータ保護法に違反する可能性もある。

(181) 評価または分類が差別から保護される根拠のいずれかに基づく場合（たとえば、年齢、宗教、人種または民族的出自、性別など）、または直接的または間接的にそれらの集団の差別をもたらす場合、そのような行為はEUの反差別法の対象にもなる。

(182) 消費者信用指令(EU)2023/2225[127]もまた、この文脈に関係し得る。消費者信用指令第18条第3項は、信用力の評価が、消費者信用の性質、期間、価値およびリスクに必要かつ相応な消費者の収入および支出ならびにその他の財務状況および経済状況に関する、関連性がありかつ正確な情報に基づき行われることを要求する。その情報には、収入もしくはその他の返済資金源の証拠、金融資産および負債に関する情報、またはその他の財務上の約定に関する情報が含まれ得る。消費者信用指令は、特別なカテゴリーの個人データが情報に含まれること、およびソーシャルネットワークから情報を取得することを特に禁止する。融資開始およびモニタリングに関する欧州銀行当局のガイドライン[128]は、信用力評価に関係する情報をさらに規定する。特別の評価を目的とするこれらの分野の法におけるこのデータの種類の仕様は、行為がAI法第5条第1項(c)の禁止の範囲内にあるかどうかを判断する際に考慮すべき関係事項である。

(183) 同様に、マネーロンダリングおよびテロリズム資金供与対策のため人の評価および分類に使用されるAIシステムは、これらの問題に関するEU法も遵守しなければならない。

5. AI法第5条第1項(d) – 刑事犯罪の個別のリスクの評価および予測

(184) AI法第5条第1項(d)は、AIシステムが、人格的特徴および特性のプロファイリングまたは評価のみに基づいて、自然人が犯罪を犯すリスクを評価しまたは予測することを禁止する。

(185) AIシステムが犯罪活動における人の関与の人による評価をサポートするために使用され、それが既に犯罪活動と直接結びつく客観的および検証可能な事実に基づく場合、この規定は、その最後の文において、その禁止事項が適用されないことを示す。プロファイリングだけでなく、人格的特徴および特性または過去の犯罪行動の評価に基づき、自然人が犯罪を犯すリスクまたは再犯を犯すリスクを評価するために、法執行機関もしくはその名で、または法執行機関を支援するEUの機関、組織、部署もしくは事務所が使用することを意図する禁止事項の適用範囲外となるAIシステムは、「ハイリスク」AIシステムに分類され（AI法附属書III、6、(d)）、かつAI法に基づくすべての関連する要件および義務を遵守しなければならない。

5.1. 理論的根拠および目的

(186) AI法前文42項は、AI法第5条第1項(d)の禁止事項の背景と理論的根拠、つまり、自然人は、それらの実際の行動に基づいて判断されるべきであり、プロファイリング、人格的特徴または特性のみに基づき、AIが予測した行動により判断されるべきではないと説明する。

---

[127] 消費者信用契約に関するならびに理事会指令87/102/EECを廃止する欧州議会および欧州理事会指令2008/48/EC、OJ L 133、2008年5月22日、66-92頁。
[128] 欧州銀行当局、2020年5月29日以降の融資開始とモニタリングに関するガイドライン、EBA/GL/2020/06

5.2. 禁止事項の主な概念と構成要素

> **AI 法第 5 条第 1 項(d)は、次のように規定する。**
>
> AI に関する以下の行為は、禁止される：
> d)自然人のプロファイリングまたはその人格的特徴または特性の評価のみに基づいて、自然人が刑事犯罪を犯すリスクを評価または予測するために、AI システムを上市し、この特定の目的のためにサービスを開始し、または使用すること；この禁止は、既に犯罪行為に直接結びつく客観的かつ検証可能な事実に基づき、犯罪行為における人の関与について、人による評価をサポート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 条第 1 項(d)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない。

(i) 当該行為は、AI システムの「上市」、「この特定の目的のためのサービスの開始」、または「使用」を構成すること。

(ii) AI システムは、自然人が刑事犯罪を犯すリスクを評価しまたは予測するリスク評価を行うものであること。

(iii) リスク評価または予測は、次の一方または双方のみに基づくものであること：

(a) 自然人のプロファイリング、

(b) 自然人の人格的特徴または特性を評価すること。

(188) 禁止事項が適用されるためには、3 つの要件すべてが同時に満たさなければならない。第 1 の要件、すなわち、AI システムの上市、サービス開始または使用は、2.3 において既に分析した。したがって、当該禁止事項は、AI システムの提供者および導入者それぞれが、その責任の範囲内において、この特定の目的のためにそのような AI システムを上市し、サービスを開始しまたは使用しないよう AI システムの提供者および導入者の双方に適用される。以下、禁止事項が適用される他の 2 つの要件を分析する。

5.2.1. 人が犯罪を犯すリスクの評価または可能性の予測

(189) 個人が犯罪を犯すリスクを評価しまたは予測するリスク評価は、個別「犯罪予測」または「犯罪予報」とよくいわれる。「犯罪予測」または「犯罪予報」の一般的に合意された定義は存在しないが[129]、これらの用語は、通常、大量の過去のデータ（社会経済データ、警察記録などを含む）に適用される各種の高度な AI 技術や分析的手法をいい、それは犯罪学理論との組み合わせによって、犯罪と戦い、管理し、および予防するために、警察および法執行機関に戦略や行動を知らせるための基礎として、犯罪を予測するために用いられる。[130]

---

[129] たとえば、オランダの犯罪認識システム(CAS)ならびにドイツおよびスイスの Precobs、Handbook のような、EU 基本権機関のハンドブック記載のシステムを参照。ハンドブック 2018 年、p.138。Preventing unlawful profiling today and in the future: a guide, Handbook, 2018, p.138.

[130] Europol, AI and policing The benefits and challenges of artificial intelligence for law enforcement, An Observatory Report from the Europol Innovation Lab, 23 September 2024 を参照。F. Yang, 'Predictive Policing' in Oxford Research Encyclopedia, Criminology and Criminal Justice, Oxford University Press, 2019 も参照。

(190)　犯罪予測 AI システムは、過去のデータ内におけるパターンを識別し、指標を犯罪発生の可能性に関連づけ、かつ予測の出力としてリスクスコアを生成する。たとえば、このようなシステムは、警察のタスクフォースのプランニングのため、ハイリスクな状況のモニタリングのため、および（再）犯罪者である可能性が予測される人の取り締まり実施のために使用され得る。このようなシステムは、法執行機関、特にリソースが不足している機関の効率を高め、犯罪を探知、防止、予測するための積極的なアプローチを可能とする好機となる。[131] しかし、他人の将来の行動を予測するために行われた犯罪に関する過去のデータをこのように使用することは、バイアスが永続化しまたは強化すらされる可能性があり、かつこれらの状況がデータセットの一部でなかったりまたは特定の AI システムが動作するアルゴリズムにおいて考慮されていなかったりする場合、重要な個別的状況が「見落とされる」結果となる可能性がある。これは、法執行および司法制度全般に対する公衆の信頼を損なう可能性もある[132]。

(191)　このようなリスク評価および予測は、原則として、将来を見据えた、かつ、将来の犯罪（まだ犯されていない）または現時点で犯されるリスクがあると評価される犯罪に関するものであり、これには、犯罪の未遂または予備行為を含む。[133] これらは、犯罪の防止や探知など、法執行活動のいかなる段階でもなされ得るだけでなく、捜査、訴追および刑事罰の執行の段階においても（司法当局が再犯のリスクを評価する場合、たとえば公判前勾留を課すことに関する決定を下す状況において）、および刑罰に服した後の社会復帰のための個別的計画の一部でもなされ得る[134]。

(192)　AI 法第 5 条第 1 項(d)の禁止事項は、犯罪予測およびリスク評価行為自体を違法とするものではない。これは、自然人が刑事犯罪を犯すリスクを評価しまたは予測するためのリスク評価を行う AI システムにのみ適用され、その場合、上記の 3 つめの要件も満たされる。さらに、前述のとおり、当該禁止事項は、AI 法第 5 条第 1 項(d)の最後の文にある明示の除外にいう状況においては適用されない。

5.2.2. 自然人のプロファイリングまたはその人格的特徴および特性の評価のみに基づくこと

(193)　AI 法第 5 条第 1 項(d)の禁止事項が適用される第 3 の要件は、自然人が犯罪を犯すリスクを評価しまたは予測するためのリスク評価は、a)当該人物のプロファイリング、または b)その人格的特徴および特性の評価のみに基づくものであることである。

---

[131] たとえば、OxRec（オランダ保護観察局、「Reclassering Nederland」）: Prediction of violent reoffending in prisoners and individuals on probation: a Dutch validation study (OxRec) - PMC (nih.gov)
[132] たとえば、EU 基本権事務所（2022 年 12 月 8 日）アルゴリズムにおけるバイアス- 人工知能と差別を参照。EU 基本権事務所。
[133] この点については AI 法前文 42 項参照。同項は、この点に関し、その「犯罪の可能性」および「実際のまたは潜在的な犯罪の発生」について述べ、それは過去形ではなく、現在形で使用される。
[134] 一例として、児童の性的虐待および性的搾取、ならびに児童ポルノに対抗する EU 指令 2011/93 第 24 条第 4 項は、刑事訴訟中の者、または児童の性的虐待に関連する行為で有罪判決を受けた者に対して、再犯の危険の評価を受けることを義務づける。

(194)　AI 法第 5 条第 1 項(d)の禁止事項は、犯罪を犯すリスクが予測されまたは評価されるすべての個人を保護することを目的とするため、AI システムが一人の自然人のみの人格的特徴および特性をプロファイリングしまたは評価するか、またはある自然人のグループの人格的特徴および特性を同時にプロファイリングしまたは評価するかにかかわらず、適用される。

a) 自然人のプロファイリング

(195)　AI 法第 5 条第 1 項(c)と異なり、第 5 条第 1 項(d)は、「プロファイリング」という用語を明示的に使用する。AI 法第 3 条(52)は、GDPR 第 4 条(4)の定義を参照し、この用語を定義する[135]。プロファイリングの概念には、その中核的要素の 1 つとして「一定の個人的側面を評価するため」という目的が含まれる。[136] AI 法第 5 条第 1 項(d)の文脈において、プロファイリングは、犯罪を犯す人のリスクを評価しまたは予測する目的で行われる。

(196)　いわゆるグループプロファイリングの概念[137] も、この文脈においては関係し得る。この概念は、特定のグループ、たとえば、犯罪の加害者（たとえば、テロリスト、ギャングなど）のカテゴリーについて、他の人が以前に犯した犯罪に関する履歴データに基づき構築された、説明的プロファイルの構築と適用をいう。これらのグループプロファイルは、後で他の人が同様の犯罪を犯すリスクを評価しおよび予測するために使用できる。AI システムが予測し、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ときは常に、個人のプロファイリングを構成し、AI 法第 5 条第 1 項(d)の禁止事項に該当し得る。

b) 人格的特徴および特性の評価

(197)　犯罪を犯す人のリスクを評価しまたは予測するためのリスク評価が、その人の人格的特徴および特性の評価のみに基づく場合にも、この禁止事項は適用される。このような評価または予測は、プロファイリングの概念に含まれることが多いが、GDPR 第 4 条(4)で定義されるプロファイリングが成立しない場合は、一つの代替的手段ともみなされ得る。

(198)　4.2.1.c において述べたとおり、人格的特徴および特性は、特定の自然人に関連する特性の広範なカテゴリーを構成し、そのために一般的に合意された分類法はない。AI 法前文 42 項は、「国籍、出生地、居住地、子どもの数、負債の程度、車の種類」など、犯罪を犯す人のリスクを予測するために評価し得る人格的特徴および特性の例を示す。これは単なる例示であり、限定列挙ではない。

---

[135] LED 第 3 条(4)は、AI 法第 5 条第 1 項(d)の禁止に関連するが、GDPR 第 4 条(4)と同じ方法で、プロファイリングを次のように定義する。「自然人に関連する一定の個人的側面を評価するための個人データの使用を構成する個人データのあらゆる形態の自動処理。特に、当該自然人の仕事におけるパフォーマンス、経済状況、健康、個人的な好み、興味、信頼性、行動、場所、または動きに関する側面を分析しまたは予測するためのもの」。 同じ定義は、EU の機関、組織、部署および事務所による個人データの処理に関する規則(EU)2018/1725 第 3 条(5)にも含まれる。OJ L 295、2018 年 11 月 21 日、39 頁。

[136] また、Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the Purpose of Regulation 2016/679, WP251rev.01, 6.2.2018, and endorsed by the EDPB, p. 7、29 条も参照。Fundamental Right Agency, Preventing unlawful profiling today and in the future: a guide, Handbook, 2018, p.138 も参照。

[137] グループプロファイリングについては、たとえば、Fundamental Right Agency, Preventing unlawful profiling today and in the future: a guide, Handbook, 2018, p. 21 参照。

c) 「のみ」

(199) AI 法第 5 条第 1 項(d)は、当該規定の対象となるリスク評価が、人のプロファイリングまたはその人格的特徴および特性の評価に「のみ」基づく場合にのみ禁止されることを規定する。AI 法前文 42 項が明らかにするところによれば、「のみ」は、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双方に対する適用が意図される。

(200) リスク評価が人格的特徴および特性のプロファイリングまたは評価「のみ」に基づくものでなければならないという条件は、多くの状況で満たされない可能性がある。

(201) AI 法第 5 条第 1 項(d)の最後の文から明らかなとおり、いかなる場合にも、AI システムが犯罪行為への人の関与の人による評価をサポートするために使用される状況が生じ、それは犯罪行為に直接結びつく客観的かつ検証可能な事実に既に基づいている。前文 42 項が明らかにするとおり、この文脈においては、特に、しかし必ずしも排他的でなく、既存の関係する自然人に関し合理的な疑いがある状況について考えなければならない。結局、そのような場合、通常、人による評価が行われることになり、それは通常、関連がある客観的かつ検証可能な事実に基づくことになる。

(202) しかし、他の状況もあり得、それは常にケースバイケースで評価される必要がある。一方では、「のみ」という用語の使用は、リスク評価において考慮されることになる他のさまざまな要素がありうることを開かれたままとしており、それにより、もはや人格的特徴または特性のプロファイリングまたは評価に基づいてのみなされるものではなくなる。他方、禁止の迂回を回避し、かつその有効性を確保するため、そのような他の要素は、禁止事項が適用されないという結論を正当化できるよう、それらにとって、すべて現実の、具体的かつ意味のあるものでなければならない。AI 法第 5 条第 1 項(d)の禁止事項をその最後の文に含まれる除外と合わせ解釈すると、特に、一定の事前に確立された客観的かつ検証可能な事実の存在が、その結論を正当化し得ることを示唆している。

たとえば、

- 法執行機関は、AI システムを使用し、個人の年齢、国籍、住所、車の種類、婚姻状況のみに基づき、テロリズムなどの犯罪について、犯罪行為を予測する。このシステムは、個人の人格的特徴のみに基づき、まだ犯していない犯罪を、個人が将来犯す可能性がより高いとみなす。このようなシステムは、AI 法第 5 条第 1 項(d)に基づき禁止されると考えることができる。

- 国税当局は、AI 予測ツールを使用し、すべての納税者の納税申告書を再検討し、潜在的な税務上の犯罪を予測し、さらなる調査が必要なケースを識別する。これは、AI システムによって構築されたプロファイルのみに基づいて行われ、それは、たとえば二重国籍、出生地、子どもの数などの人格的特徴、および特に、予測であり、それゆえに非客観的かつ検証が困難な推論された情報などの不透明な変数を、その評価に用いる。このようなシステムは、特定の人の犯罪行為への関与に合理的な疑いも存在せず、またはそれをその犯罪行為に関連づけるその他の

客観的かつ検証可能な事実もないため、通常、AI 法第 5 条第 1 項(d)の禁止事項に該当する。これは、無関係な社会的文脈から、データに不利益な取扱いを関係させる、AI 法第 5 条第 1 項(c)に基づき禁止されるソーシャルスコアリングの適用範囲に含まれる例でもある。

- 警察署は、AI ベースのリスク評価ツールを使用し、子どもや青少年が「将来の暴力的および財産的犯罪」に関与するリスクを評価する。当該システムは、他の人との関係およびその想定されるリスクの程度に基づき、子どもを評価する。これにより、子どもは、兄弟や友人など、ハイリスクと評価される他の個人に関連づけられるだけで、犯罪のリスクがより高いとみなされ得る。親のリスクの程度もまた、子どものリスクの程度に影響し得る。リスク評価の結果、警察がこれらの子どもをそのシステムに「登録」し、追加的な検査でそれらをモニタリングし、若年「ケア」サービスにそれらを紹介する。このようなシステムは、AI 法第 5 条第 1 項(d)に基づく禁止事項にも該当する可能性がある。

5.2.3. 犯罪行為に直接結びつく客観的かつ検証可能な事実に基づく人による評価をサポートするための AI システムの除外

(203) AI 法第 5 条第 1 項(d)は、その最後の文において、犯罪行為における人の関与について人による評価をサポートするために使用される AI システムに禁止事項が適用されないことを規定するが、これは既に犯罪行為に直接結びつく客観的かつ検証可能な事実に基づいている。ただし、前述のとおり、この明示の除外において規定される状況は、必ずしも禁止事項が適用されない唯一の状況ではない。禁止事項の適用範囲を示すことにより、かつ、その状況が問題となっている場合に禁止事項はいかなる場合にも適用されないことを明確にすることにより、当該条項にその状況を明示的に含めることは、法的確実性を提供することになる。

(204) システムが除外の適用範囲内にあり、それにより禁止されない場合、法執行機関またはその名で使用されることが意図され、したがって要件および人間による管理（AI 法第 14 条および第 26 条)を含む保護措置の対象となる場合、それはハイリスク AI システム（AI 法の付属書 III、6、(d)に規定されるとおり）に分類される。これらの要件は、AI システムの性能および制限を適切に理解することができ、その出力を正しく解釈し、自動化バイアスのリスクに対処できる必要な能力があり、トレーニングを受けおよび権限を有する者に、人間による管理を委ねなければならないことを含む。これらの者は、AI システムの出力を意味があるように評価するため、明確な手順、トレーニング、ならびに必要な能力および権限を有しなければならない。この特定のケースでは、人が犯罪を犯すリスクを AI が予測しまたは評価することが、犯罪行為に関連する客観的かつ検証可能な事実に基づくものであることについて、人による評価を確保しなければならない。これらの者はまた、悪影響もしくはリスクを回避するために介入するか、または AI システムが意図したように実行されない場合はその使用を停止しなければならない。

(205) さらに、「人間の介入」の概念は、欧州司法裁判所の判例の対象となってきたが、それは、特に重大犯罪に関与する航空機の乗客のリスクを予測する、自動化された意思決定のみの文脈においてである。この判例は、AI 法第 5 条第 1 項(d)において使用される「人による評価」の概念の適用にも関連し得る。

Ligue des droits humains 事件において[138]、欧州司法裁判所は、テロリズムおよびその他の重大犯罪に関与する可能性を評価するために、航空旅客の乗客名簿（PNR）データを体系的に処理する高度な AI システムの使用の合法性を検討した。

欧州司法裁判所は、指令(EU)2016/681（「PNR 指令」）のルールを、自動処理のみに基づく不利な法的判断を禁止するものと解釈し、さらに、誤った肯定的回答を識別しかつ非差別的な結果を確保するために、すべての肯定的な合致について、自動化されていない手段による個々の人による評価および検証を要求した。

欧州司法裁判所によれば、いかなる PNR データの自動処理の結果も人による評価に基づくことを条件として、人による評価は、肯定的回答が、この特定のケースにおいて、テロリスト犯罪または重大犯罪に関与している可能性のある者に関係するかどうかを評価するため、かつ自動処理の非差別的な性質を確保するため、客観的な基準に基づくものでなければならない。

(206) 除外の内容についていえば、その中心的要素の一つは、禁止事項の対象となる状況において生じるリスク評価を AI システム自体が行うことに関連するというより、AI システムが人による評価をサポートするために使用されるということである。ただし、除外が適用されるためには、当該人による評価は、加えて、犯罪行為に直接結びつく客観的かつ検証可能な事実に既に基づいたものでなければならない。

5.2.4. 民間の行為者の活動が適用範囲に入る限界

(207) AI 犯罪予測システムの主な導入者は、原則として法執行機関であるが、これに加え、民間の主体の活動も、場合によっては AI 法第 5 条第 1 項(d)の禁止事項の対象となる場合がある。その文言によれば、これは禁止事項が法執行機関だけに適用されるわけ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、禁止事項は簡単に迂回され、その有効性に疑義を生じることになる。

(208) そうであるならば、禁止事項は、特に、刑事犯罪の防止、捜査、探知もしくは訴追、または刑事罰の執行のために、民間の行為者が公的権限および公権力を行使することを法律によって委託されている場合に適用されると考えることができる[139]。また、民間の行為者は、法執行機関の名で行動すること、および個別の犯罪リスクの予測を行うことを、ケースバイケースで明示的に要請されることがある。そのような場合、適用される要件を満たし、かつ除外が適用されない場合、それらの民間の行為者の行為も禁止事項の範囲に含まれ得る。

たとえば、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は、法執行機関から、たとえば、国民の登録、銀行取引、通信データ、位置データなどの複数のソースおよびデータベースからの大量のデータを分析することや、人身売買犯罪の犯罪者となる可能性がある個人

[138] 司法裁判所 2022 年 6 月 21 日判決、Ligue des droits humains、C-817/19、ECLI:EU:C:2022:491
[139] AI 法第 3 条（45）の法執行機関の定義を参照。

のリスクを予測しまたは評価することを求められることがある。第 5 条第 1 項(d)のすべての基準を満たす場合、そのようなユースケースは禁止される可能性がある。

(209) さらに、この禁止事項は、犯罪を犯す人のリスクを評価しまたは予測する民間の主体に適用され得るが、そこでは、人が特定の犯罪を犯すリスクを評価しまたは予測する民間事業者が服する法的義務の遵守が、客観的に必要である（たとえば、マネーロンダリング防止、 テロリズム資金供与の場合)。

たとえば、銀行は、EU 反マネーロンダリング法[140]に基づき、マネーロンダリング犯罪のため顧客をスクリーニングし、プロファイリングする義務がある。銀行がその義務を履行するために AI システムを使用する場合、被疑者として特定された人物が合理的に反マネーロンダリング犯罪を犯す可能性があることを確保するため、それは、客観的かつ検証可能な当該法において特定されるデータのみに基づいて実施されなければならない。予測は、[141] また、そのような評価の正確性および適切性を確保するため、その法に従い人による評価および検証の対象としなければならない。この法を遵守することにより、反マネーロンダリング目的での個別の犯罪予測 AI システムの使用は、AI 法第 5 条第 1 項(d)の禁止事項の範囲外となることが確保される。

(210) しかし、禁止事項の文言から明らかとなる刑事犯罪の遂行に具体的かつ排他的に関係するリスク評価に焦点を当てること、および前文 42 項において説明される禁止の目的に関し、民間の主体が、特定の刑事犯罪を犯す顧客のリスクを評価しまたは予測する目的でなく、通常の事業の運営および安全のために、またはその経済的利益を保護するために（たとえば、財務上の不正の検出)、顧客をプロファイリングする場合、民間の主体の行為は、AI 法第 5 条第 1 項(d)の禁止事項の適用範囲に該当するとみなされるべきではない。

(211) 言い換えれば、法により特定の法執行業務を委託され、法執行機関の名で行動しまたは上記に述べる特定の法的義務に服する民間の当事者がない場合、民間の主体が通常の業務の過程において、およびそれ自身の私的利益を保護する目的で、リスク評価を行うために AI システムを使用することは、これらのリスク評価が純粋に偶発的かつ二次的な状況としてのみ犯罪行為が行われるリスクに関連し得る事実があるにしても、禁止事項の対象とはみなされない。

5.3. 適用範囲外となるもの

5.3.1. 位置ベースもしくは地理空間的予測または場所ベースの犯罪予測

(212) 位置ベースもしくは地理空間的または場所ベースの犯罪予測は、犯罪の場所もしくは位置、またはそれらのエリアにおいて犯罪が行われる可能性に基づく。原則として、このような取締りは特定の個人の評価とは無関係である。したがって、それらは禁止の範囲外である。

位置ベースもしくは地理空間的予測または場所ベースの犯罪予測の例

[140] 2024 年 5 月 31 日のマネーロンダリング防止規則(EU)2024/1624。
[141] 規則(EU)2024/1624 第 20 条。

- AI ベースの予測取締りシステムは、エリアごとの過去の犯罪率およびストリートマップなど他の補足情報に基づき、市内のさまざまなエリアにおける犯罪の可能性のスコアを提供し、たとえば強盗、ナイフによる犯罪など、特定の種類の犯罪のリスクが高まっていることを強調し、法執行機関が犯罪行為を中断および停止させるための地区取締りを実行するために、警察のパトロール/存在を減らしまたは増やして動員する決定に役立つ。

- 税関当局は、AI リスク分析ツールを使用し、たとえば既知の密輸ルートに基づき、麻薬または禁制品の場所の可能性を予測する。

- 警察署は、AI により動作するシステムを使用し、銃声をリアルタイムで探知し、位置を特定する。当該システムは、都市部の音響センサーを使用し、銃声を識別し、その位置を三角測量し、犯罪の探知および捜査に役立つ実用的なデータを警察官に提供する。

(213) しかし、位置ベースの犯罪予測システムと、人が犯罪を犯すリスクを評価する個々の予測システムとをどのように区別するかは、必ずしも明確であるとはいえない。AI システムが位置ベースの予測取締りを実行し、そこで個人のプロファイリングの側面としてその位置のリスクスコアを考慮する限り、そのシステムは人ベースと見なされ、他の理由で禁止の範囲外になる可能性があるとしても、原則として AI 法第 5 条第 1 項(d)の対象とされるべきである。

たとえば、位置ベースのもしくは地理空間的情報、または場所ベースの情報が個人に関する情報（たとえば人の居住地）と関連づけられ、かつ、AI システムが、当該人が、犯罪が多いその居住地を含む、問題の個人のプロファイリングのみに基づいて、犯罪を犯す可能性が高いとのリスクを評価する場合、当該システムは人ベースであるとみなされなければならない。

5.3.2. 犯罪行為に結びつく客観的かつ検証可能な事実に基づき人による評価をサポートする AI システム

(214) AI 法第 5 条第 1 項(d)が規定するところによれば、既に犯罪行為に直接結びつく客観的かつ検証可能な事実に基づき、犯罪行為に対する自然人の関与の人による評価をサポートするために使用される AI システムには、禁止事項は適用されない。このような場合、個別の犯罪リスクの評価および予測はまた、プロファイリングまたは人の特徴の評価のみに基づくわけではなく、したがって禁止されない。

この理由により禁止の範囲外となる AI システムの例には、以下を含む：

- 実際の行動のプロファイリングおよび分類のための AI システムの使用。群衆の中で合理的な疑いのある危険な行動のような、誰かが準備しかつ犯罪を犯す可能性があり、そこで AI の分類により意味のある人による評価が行われること。この場合、AI のサポートで人により行われるリスク評価は、個人の特性またはプロファイリングに基づくだけでなく、行動の前に人により検証されたその者の脅迫的な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。

- 警察は武装強盗の可能性のリスクを調査するなか、2 名の個人を疑っている。たとえば、検証可能な参入、および武器購入のためのダークウェブチャットグループでの対話など、その疑惑の根拠となるいくつかの検証可能かつ客観的な事実が存在する。容疑者所有の車両の地理空間的予測または場所ベースの取締り情報および自動的ナンバープレート登録（ANPR）の情報を組み合わせる AI システムは、特定の犯罪行為に直接結びつく検証可能かつ客観的な事実に基づき捜査における人による評価をサポートする。

- 受刑者が早期釈放の恩恵を受けるべきかどうかのリスクを評価する AI システムの使用。対象者の AI プロファイル、またはその人格的特徴および特性の評価は、過去の犯罪に結びつく客観的かつ検証可能な事実およびリハビリテーションに関連する実証された行動について、人による評価をサポートするだけである。

- 裁判官は、重大な刑事犯罪で訴追された者に対し公判前勾留質問を行い、身柄拘束しない措置を適用できるかどうか評価する。当該決定は、被疑者または被告人が勾留されない場合に他の犯罪を犯す可能性や、逃亡または捜査の適切な実施を妨害する可能性など、公判前勾留を行う正当な理由の存在の評価に基づく。このプロセスを支援するため、裁判官は、同様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基づきトレーニングされた AI リスク評価ツールを使用する。

- AI システムは、身柄拘束なく刑に服する個人が、釈放条件に違反するリスクまたは逃亡するリスクを、過去の犯罪行為に基づき、および釈放条件の遵守、心理的評価の結果、個人が使用する可能性がある他のコミュニティサービスからの推奨など、容疑の根拠となる客観的事実に基づき評価するために、人の担当者の評価のサポートに使用される。この情報に基づき、担当者は、その現状を維持するか、釈放の条件を見直すかを決定する。

- EU に入る商品が国境において適用される法を遵守していないリスク（たとえば、違法薬物の輸入禁止、輸出制裁違反またはその他の違法行為を含み得る）を評価し、税関における管理を実施すべき状況を識別するために、税関当局が使用する AI システム。当該 AI システムは、商品とそのサプライチェーンに関して税関に提供された客観的かつ検証可能な情報（たとえば、商品の性質および価値、コンテナ番号、他の商品を隠匿するための輸送手段、EU への輸入または EU からの輸出の要件に商品の特定の説明および原産地が準拠していることに関する事前の認識など）を評価する。一定の場合、輸出入を行う者の、商品の輸入に関連する不正行為、犯罪組織への所属、または麻薬取引の犯罪歴に関する以前の関与の情報も処理することもある。このようなシステムは禁止行為の適用範囲外であるが、それは、自然人が違法な商品の輸出入に関与する可能性の予測が、プロファイリングだけでなく、商品および輸出入を行う者の犯罪行為への以前の関与に関する客観的かつ検証可能な情報に基づくものであり、状況が税関における管理またはリスク緩和行為を要するかどうかを判断するために人による検証の対象となるからである。

5.3.3. 法人に関係する犯罪の予測および評価のために使用される AI システム

(215) AI 法第 5 条第 1 項(d)の禁止事項は、自然人の個別の予測とリスク評価にのみ適用される。したがって、一般的に、企業または非政府組織などの法人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。

たとえば、 -税務当局または税関当局は、犯罪を構成する脱税または税関における不正行為を企業が犯すリスクを評価するために、AI システムを使用し、企業の取引、税務申告および関税データに関する大量のデータを分析する。 -違法な商品が EU に送られないよう法人に指示を出すべき状況を識別することをサポートするために、税関当局が使用する AI システム。

(216) 同時に、自然人が法人を通じ、「個人事業主」として、または独立した専門家（たとえば弁護士）として行動する、境界線上のケースもあり得る。このような状況においては、AI システムが特定の自然人を識別し、その犯罪を犯すリスクを評価しまたは予測するので、たとえそれが自然人が行う商業活動に関連する目的で行われた場合でも、すべての要件が満たされていることを条件として、AI 法第 5 条第 1 項(d)の禁止事項は適用され得る。

5.3.4. 行政犯罪の個別的予測のために使用される AI システム

(217) AI 法第 5 条第 1 項(d)の禁止事項は、刑事犯罪の予測にのみ適用され、したがって、その訴追が人の基本的権利および自由により介入的でない行政犯罪を、原則としてその適用範囲から除外する。

たとえば、軽微な犯罪（軽微な交通違反など）を犯す可能性のある犯罪者のリスクを評価するため、または税金、仕入れ、経費プロセスにおける不正を評価するため、行政上の調査の文脈において AI を使用する公的機関は、行政上の調査および確認の結果、自然人が犯罪に関与する可能性について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 条第 1 項(d)の禁止事項の適用範囲に該当しない。

(218) 犯罪が、その性質上、行政上のものか刑事上のものかは、EU 法または国内法次第である。EU 法により直接規制されていない犯罪に関し、「刑事犯罪」は EU 法のなかで自律的な意味を持つ概念であり、かつ加盟国間で一貫するように解釈されるべきであるから、犯罪の国内における性質は、欧州司法裁判所による検討対象となる。欧州司法裁判所は、別の文脈において、加盟国による犯罪の分類はその観点において最終的ではないと結論づけた[142]。犯罪の性質（犯罪者か否か）を評価するために使用される関連の基準は、欧州司法裁判所および欧州人権裁判所（ECtHR）の関連の判例法に見出され得る。[143]

---

[142] たとえば、2013 年 11 月 14 日の裁判所（大法廷）判決 - Marián Baláž に対して発せられた制裁金の執行に関する手続き Case C-60/12, ECLI identifier: ECLI:EU:C:2013:733 参照。

[143] 欧州司法裁判所の判例法によれば、国内裁判所が、いわゆる「エンゲル基準」に照らし、刑事罰でないものが「犯罪者」とみなされるかどうかを決定する、欧州人権裁判所（ECtHR）1976 年 6 月 8 日判決、Engel and Others v. the Netherlands、Application nos5100/71、5101/71、5102/71、5354/72 および 5370/72、CE:ECHR:1976:0608JUD000510071、82 項参照。元来、欧州人権裁判所（ECtHR）によって発展し、その後、欧州司法裁

5.4. 他の EU 法との相互作用

(219) AI 法第 5 条第 1 項(d)の禁止事項と LED および GDPR との相互作用は、GDPR や LED などの EU データ保護法に基づく個人データ処理の合法性を評価する場合に関連する。特に、AI 法第 5 条第 1 項(d)は、法執行機関、その他の公的機関、および禁止事項の適用範囲内にある民間の主体に対し、自然人のプロファイリングのみに基づき、または自然人の人格的特徴および特性の評価のみに基づき、自然人が犯罪を犯すリスクを評価しまたは予測することについて、特別な禁止を課す。LED に関し、AI 法第 5 条第 1 項(d)は、差別（直接的または間接的な）をもたらすプロファイリングを禁止する LED 第 11 条第 3 項を害しない。

(220) AI 法第 5 条第 1 項(d)の禁止事項と無罪の推定に関する指令(EU)2016/343 との相互作用もまた、関連する。双方の行為は、法に従い有罪が証明されるまで無罪と推定される基本的権利と、指令の場合には直接的に、AI 法の場合には間接的に（前文 42 項参照）、関連性があるからである。[144] 当該指令は、ある者が刑事犯罪を犯したと疑われ、または告発されたときから適用されるが[145]、AI 法はより広い適用範囲を有し、特定の人に対し正式な犯罪捜査が開始される前の予測および犯罪防止の段階で、たとえそのような予測およびリスク評価が AI 法第 5 条第 1 項（d)の適用を受ける民間の行為者により行われ、司法当局を含む管轄の法執行機関によるものでない場合でも、AI 法が既に適用される。

(221) AI 法第 5 条第 1 項(d)の禁止事項が適用されない場合でも、特にデータ保護法、刑事訴訟法および警察法、ならびに個々の犯罪予測 AI システムの使用をさらに制限しまたは追加的条件を課し得る保護措置を含む、適用される EU 法および国内法が、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 条第 1 項(e)- 顔画像の無差別なスクレイピング

(222) AI 法第 5 条第 1 項(e)は、インターネットまたは CCTV の映像から顔画像を無差別にスクレイピングすることにより、顔認識データベースを作成または開発する AI システムを上市し、この特定の目的のためにサービスを開始し、または使用することを禁止する。

判所により承認されたこれらの基準は、選択的なものであり、累積的なものではない。罰則が刑事の性質を有するかどうかを検討する場合、管轄の国内裁判所は、以下を判断しなければならない：(1)国内法に基づく関連規定の分類；(2)犯罪の性質自体；(3)罰則の厳しさ。犯罪の性質を評価する際に考慮される側面には、特に、以下を含む：手続きが法令による執行権限を有する公的機関により実施されるかどうか；法的なルールが懲罰的または抑止的な目的を有するかどうか；法的なルールが、通常は刑法によって保護される社会の一般的利益の保護を追求するかどうか；有罪の認定により、なんらかの罰則が科されるかどうか。罰則の厳しさに関し、関連する参照情報は、国内法で規定されている課され得る罰則の上限である。これらの基準は選択的なものであり、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 条に関するガイド、公正な裁判を受ける権利（刑事関係）、2024 年 2 月 29 日更新参照。欧州司法裁判所 2012 年 6 月 5 日判決、Bonda、Case C-489/10、EU:C:2012:319、37 項 ff；欧州司法裁判所 2013 年 2 月 26 日判決、Åkerberg Fransson、Case C-617/10、EU:C:2013:105、35 項も参照。

[144] 無罪の推定は、EU 基本権憲章第 48 条に記される基本的権利である。

[145] 欧州司法裁判所が明確にするとおり、指令を適用するために、その者が管轄当局により被疑者/被告人としてその立場を認識していることは要件ではない。

6.1. 理論的根拠および目的

(223) インターネットおよび CCTV の映像から顔画像を無差別にスクレイピングすることは、個人のプライバシーに対する権利およびデータ保護に対する権利を著しく妨げ、それらの個人が匿名のままでいる権利を否定する。したがって、AI 法前文 43 項は、「大衆監視の感覚」および「プライバシーの権利を含む、基本的権利の重大な侵害」のリスクに基づき、AI 法第 5 条第 1 項において(e)で定める禁止事項を正当化する。

6.2. 禁止事項の主な概念および構成要素

> **AI 法第 5 条第 1 項(e)は、次のように規定する。**
>
> AI に関する以下の行為は、禁止される：
> (e) インターネットまたは CCTV 映像から顔画像を無差別にスクレイピングすることにより、顔認識データベースを作成しまたは開発する AI システムを上市し、この特定の目的のためにサービスを開始し、または使用すること；

(224) AI 法第 5 条第 1 項(e)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない。

(i)当該行為は、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」を構成するものであること；

(ii) 顔認識データベースを作成または拡張する目的であること；

(iii)データベース化する手段は、無差別のスクレイピングのための AI ツールを通じて行われること；および

(iv) 画像の出所がインターネットまたは CCTV の映像のいずれかであること。

(225) 禁止事項が適用されるためには、4 つの要件すべてが同時に満たされなければならない。AI システムの上市、サービス開始または使用という第 1 の要素は、既に 2.3 において分析している。したがって、この禁止事項は、AI システムの提供者および導入者の双方に適用され、それぞれがその責任の範囲内において、そのような AI システムを上市し、サービスを開始しまたは使用してはならない。無差別のスクレイピングの禁止に関する特別の基準は、以下でさらに説明し、分析する。この禁止事項は、インターネットまたは CCTV の映像から顔画像を無差別にスクレイピングする「この特定の目的」で上市されまたはサービスが開始されるスクレイピングツールに適用される。これは、この禁止事項が、顔認識用のデータベースを構築しまたは開発するためのあらゆるスクレイピングツールに適用されるのではなく、無差別にスクレイピングするためのツールにのみ適用されることを意味する。

6.2.1. 顔認識データベース

(226) AI 法第 5 条第 1 項(e)の禁止事項は、顔認識データベースの作成または開発に使用される AI システムを対象とする。この文脈における「データベース」は、コンピュータによる迅速な検索

のために特別に構築された、データまたは情報のあらゆる集合体をいうものと理解されなければならない。顔認識データベースは、顔のデータベースに対し、デジタル画像またはビデオ映像から人間の顔を照合し、データベース内の画像と比較して、両者の間に一致の可能性があるかどうか判断することを可能とする。このような顔認識データベースは、一時的、集中的、または分散的であり得る。第5条第1項(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を要求するものではない；データベースが顔認識に使用することができるだけで十分である。

6.2.2. 顔画像の無差別なスクレイピングによる

(227) 「スクレイピング」とは、一般に、ウェブクローラー、ボット、またはその他の手段を使用し、CCTV、ウェブサイトまたはソーシャルメディアなどのさまざまなソースから、データまたはコンテンツを自動的に抽出することをいう。これらのツールは、データベースを通じて選別し、情報を抽出し、その情報を別の目的で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) 「無差別に」とは、具体的かつ個別的に意図されたスクレイピングの対象を標的とすることなく、できるだけ多くのデータおよび情報を吸収する「掃除機」のように動作する技術に関係する。スクレイピングは、データまたはコンテンツを無差別に収集する。したがって、「無差別に」という概念は、ある特定の個人または個人のグループに特定の焦点を絞らないことを意味する。robot.txt などのインターネットプロトコルのオプトアウトの尊重は、スクレイピングの無差別という性質に影響を与えない。

(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人のグループのみの人間の顔を含む画像またはビデオを収集するよう指示された場合、スクレイピングは、たとえば、特定の犯罪者を見つけたり、被害者のグループを識別したりするなど、無差別でないものとなる。このようなスクレイピングは、AI 法第5条第1項(e)の禁止事項の対象外である。

(230) たとえば、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ピックアップするために、クローラーを使用して被害者のある分類に焦点を当てた画像を無差別にではなく収集することは、禁止の対象外である。無差別なスクレイピングは、禁止行為の回避を許さないように解釈されなければならない。データベースを段階的に作成するために、インターネットまたはCCTV の映像をスクレイピングし、それによってその都度、特定の個人のグループやその他の基準を選択することは、最終的な結果は最初から無差別なスクレイピングを行うことと機能的に同じであり、AI 法第5条第1項(e)の禁止事項に含まれなければならない。

(231) システムが画像または動画の無差別でない検索と無差別な検索とを組み合わせる場合、無差別なスクレイピングは禁止される。

6.2.3. インターネットおよびCCTV の映像から

(232) AI 法第 5 条第 1 項(e)の禁止事項が適用される場合、顔画像の出所は、インターネットまたは CCTV の映像のいずれかとなり得る。インターネットに関しては、ある人が自分の顔画像をソーシャルメディアプラットフォームに公開している事実は、その人が顔認識データベースにそれらの画像を含めることに同意したことを意味するものではない。CCTV の映像から顔画像をスクレイピングする例として、空港、道路、公園などの場所において運用される監視カメラによって得られた画像が含まれる。

例：

顔認識ソフトウェア会社は、顔写真を収集する。同社が保有する写真は、インターネットを検索して人間の顔を含む画像を検出する「自動画像スクレイパー」を使用し、ソーシャルメディア（たとえば、Facebook、YouTube、Twitter、Venmo)からスクレイピングされる。これらの画像は、関連する情報（たとえば、画像のソース(URL)、ジオローカライゼーション、ときには個人の名前）とともに収集される。そして、顔の特徴が画像から抽出され、数学的表現に変換され、インデックス化および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、その画像がデータベース内の顔と一致するかどうかをシステムが判断する。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く、数学的変換が行われる。

(233) AI システムが人物の写真を受け取り、一致させるためにインターネット上で顔を検索する、すなわち「リバースエンジニアリング画像検索エンジン」の場合、これは無差別でないスクレイピングとみなされる。その上、一致が「データベース」に示されるかどうかは疑問である。

6.3. 適用範囲外となるもの

(234) AI 法第 5 条第 1 項(e)の禁止事項は、顔画像以外（音声サンプルなど）の生体データの無差別のスクレイピングには適用されない。AI システムがスクレイピングに関係していない場合も、この禁止事項は適用されない。また、人物が識別されない、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。

(235) AI 法第 5 条第 1 項(e)の禁止事項は、架空の人の新たな画像を生成する AI モデルを構築するためにインターネットから大量の顔画像を収集する AI システムには適用されない。このようなシステムは、実際の人を認識する結果にならないからである。このような 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 条第 1 項(e)の禁止事項は、顔認識データベースの作成または開発に使用される AI システムを対象とする。禁止事項の適用開始前に構築された既存の顔データベースが、AI 対応の無差別なスクレイピングによるさらなる開発がない場合、それらのデータベースおよびその使用は、適用される EU データ保護ルールを遵守しなければならない。

(237) AI 法第 5 条第 1 項(e)の禁止事項は、顔認識データベースの作成または開発を対象とする。生体識別の具体的な行為は、AI 法およびその他の関連する EU 法の特定のルールの対象となる。

6.4. 他の EU 法との相互作用

(238) EU データ保護法に関連し、顔認識データベースを構築または開発するためのインターネットまたは CCTV 素材の無差別なスクレイピング、すなわち個人データ処理（データの収集およびデータベースの使用）は違法であり、かつ GDPR、EUDPR および LED に基づく何らの法的根拠にも依拠し得ない。

7. AI 法第 5 条第 1 項(ｆ) 感情認識

(239) AI 法第 5 条第 1 項(f)は、AI システムの使用が医療上または安全上の理由による場合を除き、職場および教育機関の領域において、AI システムが自然人の感情を推測することを禁止する。AI 法附属書 III(1)(c)により、禁止事項に該当しない感情認識システムは、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、一定の透明性要件を定める。

7.1. 理論的根拠および目的

(240) 感情認識技術は急速に発展しつつあり、人から感情を検出し、収集し、分析し、分類し、反応し、協働し、学習するための、さまざまな技術および処理操作に及ぶ。このような技術は、「感情技術（affect technology）」ともいう。感情認識は、適用範囲が幅広く[146]、複数の領域および分野において使用され得る。たとえば、顧客行動の分析[147]、およびターゲティング広告、ニューロマーケティング[148]のため；エンターテインメント業界において、たとえば、パーソナライズされた推奨を提供するため、または映画への反応予測のため；医療およびヘルスケアにおいて、たとえば、うつ病の検出のため、自殺予防のため、自閉症の検出のため、または教育において、たとえば、学習者（異なる年齢の生徒や学生）の注意または関与を監視するため；雇用において、たとえば、採用プロセスに加えたり、従業員の感情や退屈だけでなく、「労働者をより幸せにする」[149]ためのウェルビーイング適合性を監視するため；たとえば、嘘発見器または大きなイベントにおける感情スクリーニングなど、法執行機関や公共の安全のため;その他多くの目的のためである。

---

[146] 経済的な目的で感情を利用することは、「エモーショノミクス（emotionomics）」ともいう。
[147] たとえば、G. Mangano, A. Ferrari, C. Rafale, E. Vezzetti, F. Marcolin, 'Willingness of sharing facial data for emotion recognition: a case study in the insurance market' in AI & Society, London, Springer, 2023 参照.
[148] N. Lee, A. J. Broderick, & L. Chamberlain, 'What is 'neuromarketing'? A discussion and agenda for future research' International Journal of Psychophysiology、63(2)、2007,199-204 参照。これによれば、ニューロマーケティングは、研究分野として、「市場およびマーケティング・エクスチェンジに関する人間の行動を分析しおよび理解するための神経科学手法の適用」と定義される(200 頁)。
[149] E. Ackerman, & E. Strickland, 'Are you Ready for Workplace Brain Scanning? Extracting and using brain data will make workers happier and more productive, backers say', IEEE Spectrum, 19 November 2022, https://spectrum.ieee.org/neurotech-workplace-innereye-emotiv 参照。著者は、次のように説明する。「センサーは脳のさまざまな領域にわたる電気的活動を検出し、かつ、その活動におけるパターンは、ストレス、集中力、外部刺激への反応など、さまざまな感情や生理学的反応と広い相関関係があり得る」。

(241) 感情認識は、その有効性または正確性に関し、よく疑われる[150]。AI 法前文 44 項が説明するとおり、「感情を識別しまたは推論することを目的とする AI システムの科学的根拠については、感情表現が文化や状況によって大きく異なり、同じ個人においてすら大きく異なるだけにいっそう深刻な懸念がある。当該システムの主要な欠点は、とりわけ、その限定的な信頼性、その正確性の欠如、およびその限定的な一般的適用可能性である」。それは、さらに、感情認識は、「差別的な結果をもたらし、かつ関係者の権利および自由に介入的である可能性がある」と説明する。特に、プライバシー、人間の尊厳、思想の自由に対する権利を侵害する可能性がある。これは、特に、職場ならびに教育機関および職業訓練機関の文脈における、労働者や学生の双方が特に脆弱な立場にある非対称の関係において、重要な役割を演じる。同時に、安全や医療（たとえば治療や診断など）のためなど、特定の使用の文脈における感情認識は、利点がある。[151]

7.2. 禁止事項の主な概念および構成要素

> AI 法第 5 条第 1 項(f)は、次のように規定する。
> AI に関する以下の行為は、禁止される：
> f)職場および教育機関内において、自然人の感情を推測するために、AI システムを上市し、この特定の目的のためにサービスを開始し、または使用すること。AI システムの使用が、医療上または安全上の理由により、導入されまたは上市されることを目的とする場合を除く；

(242) AI 法第 5 条第 1 項(f)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない：

(i) 当該行為は、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 AI システムの「使用」を構成すること；
(ii) 感情を推測する AI システムであること[152]；
(iii) 職場または教育および訓練機関内であること；および
(iv) 医療上または安全上の理由による AI システムは、禁止事項から除外される。

(243) 禁止事項が適用されるためには、4 つの要件すべてが同時に満たされなければならない。第 1 の要素、すなわち、AI システムの上市、サービス開始または使用は、2.3 において既に分析した。したがって、この禁止事項は、AI システムの提供者および導入者のそれぞれがその責任の範囲内において、そのような AI システムを上市し、サービスを開始し、または使用しないよう、双方に適用される。禁止に関するその他の要件は、以下でさらに説明し、分析する。

7.2.1. 感情を推測するための AI システム

a) 感情を推測するための AI システム

[150] たとえば、J. Stanley, Experts Say 'Emotion Recognition' lacks Scientific Foundation, 18.7.2019, ACLU, referring to a study by L. Feldman Barrett e.a., 'Emotional Expressions Reconsidered: Challenges to Infering Emotion From Human Facial Movements', Psychological Science in the Public Interest, 2019, pp.iii-90.参照
[151] たとえば、R. El Kaliouby and R. Picard and S. Baron-Cohen, 'Affective Computing and Autism', Annals New York Academy of Sciences, 2007, pp.228-248 参照。
[152] または、テクノロジーが感情を推測できること（すなわち上市の場合）。

(244)　AI 法第 3 条(39)は、「感情認識システム」を「自然人の生体データに基づき、その自然人の感情または意図を識別または推測することを目的とする」AI システムと定義する。AI 法第 5 条第 1 項(f)の禁止事項は 、「感情認識システム」ではなく、「自然人の感情を推測する AI システム」のみである。前文 44 項は、さらに、その禁止事項が「感情を識別または推測する」AI システムを対象とすることを明確にする。

(245)　推論は、一般に、前提条件として識別を包含するため、禁止事項は、感情または意図を識別するまたは推論する AI システム双方を含むと理解されなければならない。[153] 一貫性の理由により、AI 法第 5 条第 1 項(f)の禁止事項は、他の感情認識システムに適用されるルール（付属書 III（1）(c)および AI 法第 50 条)と同じ適用範囲を有すると解釈すること、および人の生体データに基づく推論にそれを限定することも重要である。したがって、AI 法第 3 条（39）における感情認識システムの定義は、AI 法第 5 条第 1 項(f)との関係において考慮されなければならない。

b）感情または意図の識別および推論

(246)　「識別」は、自然人の生体データ（たとえば、声や顔の表情）処理により、感情認識システムにおいて、以前にプログラムされた感情とある感情を直接的に比較し、かつ識別することができる場合に生じる。「推論」は、分析プロセスおよびその他のプロセスにより生成された情報をシステム自体によって推測することによってなされる。このような場合、感情に関する情報は、自然人について収集されたデータに基づくだけではなく他のデータから推測され、それにはデータから感情を探知する方法を学習する機械学習アプローチなどを含む。[154]

c）感情

(247)　AI 法第 5 条第 1 項(f)の目的のため、感情または意図の概念は、広い意味で理解され、限定的に解釈されるべきではない。AI 法前文 18 項は、「幸せ、悲しみ、怒り、驚き、嫌悪、困惑、興奮、恥かしさ、軽蔑、満足、楽しさなど」の感情を列挙し、いくつかの詳細を提供する。これらの例は、限定列挙ではない。

(248)　この禁止事項は、態度を参照することにより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、たとえば人が怒りの態度を示していることを見出す場合を含む。

(249)　AI 法前文 18 項は、「たとえば、事故防止の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、痛みや疲労のような、身体的状態」を、感情または意図に含まないことを明確にする。さらに、それは、「直ちにはっきりする表情、しぐさまたは動作の単なる検知も、それらが感情の識別や推測に使用される場合を除き」、対象としないことを明確にするが、これは AI 法第 5 条第 1 項(f)にも適用されると理解されなければならない。これら

[153] AI 法前文 18 項も参照。
[154] AI 法前文 12 項参照。したがって、推論されたデータは、データセットにおける相関関係を見いだすこと、およびパターンを見いだすことを目的とする確率ベースの分析（ビッグデータ）プロセスの結果であることもよくある。

の表現としては、眉をひそめたり微笑んだりするなどの単なる顔の表現、または手、腕、頭の動きなどのしぐさ、さらには大声を上げたりささやいたりするなどの人の声の特徴もあり得る。ただし、これらの直ちに明らかな表現やジェスチャーが感情または意図を識別しまたは推測するために使用される場合、それらは禁止の対象となる。

たとえば - 人が笑っているという観察は感情認識ではない。 - 人が病気であるかどうかを識別することは、感情認識ではない。 -ニュースキャスターがカメラに向かって何回微笑んだかを追跡できるデバイスをテレビ放送局が使用することは、感情認識ではない。 - 人が幸せであると結論づけることは、感情認識である。従業員が、顧客に対し、(身振り、眉をひそめる、笑顔の欠如などから) 不満、不機嫌または怒っていると推測する AI システムは、「感情認識」である。 -学生が激怒して暴力的になりそうだと声や身振りから推測するシステムは、「感情認識」である。 -プロのパイロットまたはドライバーの疲労を推測し、それらに警告し、事故を避けるべくいつ休憩をとるか提案するために、AI 認識システムを使用することは、「感情認識」ではない。感情認識には痛みや疲労などの身体的状態が含まれていないためである。

d) その生体データに基づく

(250) AI 法第 3 条 (39) の定義によれば、生体データに基づき感情または意図を識別しまたは推測する AI システムのみが感情認識システムを構成する。[155]

(251) 生体データが抽出され得る個人的特性は、身体上または行動上の属性である。生理学的生体は、指紋、虹彩のパターン、顔の輪郭、手の静脈の形状など、人の身体的、構造的および比較的静的な属性を用いる。いくつかの様相は本質的に微視的であるが、DNA や臭気など、やはり取得されおよび識別され得る生物学的および化学的構造を示す。[156]行動上の生体は、あるタスクまたは一連のタスクを実行する際の個人の動作、身振りおよび運動能力の区別可能な特性を監視する。これは、歩行 (歩き方の分析) やキーボードとの指の接触 (キーストローク) などの人間の動作がキャプチャされ、分析されることを意味する。行動上の生体は、自発的および自発的でない反復運動、および署名、歩き方、声、キーストロークから、アイトラッキングおよび心拍[157]、脳波

[155] AI 法第 3 条(34):「生体データ」を「顔画像や指紋データのような、自然人の身体的、生理的、または行動的特徴に関連する、特有の技術的処理の結果である個人データ」と定義する。声および話し方からの感情推論について、AI 法前文 18 項も参照
[156] Physiological and Behavioural Biometrics - Biometrics Institute
[157] Physiological and Behavioural Biometrics - Biometrics Institute

計(EEG)、[158]心電図(ECG[159])に至るまで、身体の特徴に関連するリズミカルなタイミング/圧力の双方を示す、さまざまな様相を含む。生体入力は、1 つの様相（たとえば顔画像)または複数の様相（たとえば脳波図(EEG)と組み合わせた顔情報）に関連づけることができる。前文 18 項は、顔の表現、手の動きなどのしぐさ、人の声の特徴を例として示す。

たとえば、 -一定の記事のスタイルやトーンを定めるため、記述されたテキストから感情を推測する AI システム（コンテンツ/センチメント分析）は、生体データに基づいていないため、禁止事項の範囲に入らない。 -キーストローク（タイピングの仕方)、顔の表現、体の体勢または動きから感情を推測する AI システムは、生体データに基づいており、禁止の範囲に入る。

(252) したがって、AI 法による生体データの定義は広範であり、感情認識、生体分類または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. 職場および教育機関に対してという禁止事項の限定

(253) AI 法第 5 条第 1 項(f)の禁止事項は、「職場および教育機関内」における感情認識システムに限定される。AI 法前文 44 項で明確にされるとおり、この制限は、仕事または教育の文脈における権力の不均衡に取り組むことを意図する。

a) 「職場」

(254) 「職場」の概念は広く解釈されるべきである。この概念は、自然人がその雇用主によって、または、たとえば、自営業の場合、それらが所属する組織によって、割り当てられた業務および責任に従事する特定の物理的空間または仮想空間に関連する。これには、仕事が行われるあらゆる環境を含み、屋内のオフィススペース、工場、倉庫から、店舗、スタジアムまたは博物館などの公衆がアクセスできるスペース、屋外のサイトまたは車、および一時的または移動可能な作業現場までにわたり、仕事の性質に基づき多様となり得る。これは、従業員、受託者、研修生、ボランティアなどの地位に関係ない。[161] AI 法第 5 条第 1 項(f)における「職場」の概念は、選考および採用プロセス中にある候補者に適用されると理解されるべきであり、それは、雇用、労働者管理、および自営業へのアクセスの分野において、AI システムを上市し、サービスを開始しまたは

[158] EDPS 参照。TechDispatch 1/2024 – Neurodata, 3.6.2024 そこでは、脳データの使用と関連技術が論じられ、および精神的なプライバシーおよび完全性を含む新しい「ニューロライツ」の提案を含む、法的意味について論じる。S. O'Sullivan, H. Chneiweiss, A. Pierucci and K. Rommelfanger, Neurotechnologies and Human Rights Framework: Do we need new Human Rights?, Report, OECD and CoE, 9.11.2021, p.33 では、ニューロテックの最先端および法的側面について論じる。

[159] Hasnul et al., 2021, Electrocardiogram-Based Emotion Recognition Systems and Their Applications in Healthcare 参照。

[160] AI 法においては、GDPR の生体データの定義と異なり、生体データの定義には「一の識別を許可しまたは確認する」（生体データの機能的使用)という文言は含まれないが、GDPR においてはこの要件を含む。（さらに、たとえば、GDPR 第 9 条第 1 項および第 9 条第 2 項が適用される場合)、GDPR による生体データの定義は、個人データ処理に関するデータ保護ルールの下で適用されることになる。

[161] また、職場におけるハイリスク AI システムに関連する前文も参照。たとえば、前文 56 項。これは、広範な解釈を展開する。附属書 III におけるハイリスク AI システムのリストも参照。これは、4 で自営業について述べる。自営業は、EU の非差別法によっても広く対象とされる。

使用することを扱うAI法の他の規定と一致する。これは、力の不均衡があること、および感情認識の介入的な性質が採用段階で既に適用され得るからである。

たとえば：

- 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、従業員の怒りなどの感情を追跡することは禁止される。[162]個人的トレーニングの目的でのみ導入される場合、感情認識システムは、その結果が人事責任者と共有されず、トレーニングを受けた人の評価、昇進などに影響を与え得ない場合は、許可される。ただし、禁止事項が回避されることなく、感情認識システムの使用が仕事上の関係に何ら影響を与えないことを条件とする。

- コールセンターが音声認識システムを使用し、怒りや苛立ちなど顧客の感情を追跡することは、AI法第5条第1項(f)により禁止されない（たとえば、従業員が怒っている一定の顧客に対応することを支援するため)。

-ハイブリッドなワークチームにおいて、ハイブリッドビデオ通話の音声や画像から感情を識別し推測することにより、感情のトーンをモニタリングするAIシステムは、通常、社会的認識、感情力学管理、および紛争予防を促進する目的に資するものであるが、禁止される。

- 採用プロセスにおける感情認識AIシステムの使用は禁止される。

- 試用期間中の感情認識AIシステムの使用は禁止される。

- スーパーマーケットがカメラを使用し、従業員の幸せなどの感情を追跡することは禁止される。

- スーパーマーケットや銀行が不審な顧客を探知するためにカメラを使用すること、たとえば誰かが強盗しようとしていると結論づけることは、何ら従業員が追跡されないこと、および十分な保護措置が講じられていることが確認される場合、AI法第5条第1(f)による禁止の対象ではない。

b）「教育機関」

（255）教育機関の意味は広範であり、公立および私立の機関の双方を含むと理解されなければならない。生徒もしくは学生の種類もしくは年齢、または特定の環境（オンライン、対面、混合方式など）に関して制限はない[163]。たとえば、あらゆるレベルの教育および訓練機関が、AI法第5条第1項(f)の禁止事項の対象となり、これには、職業訓練校、すなわち、生徒がその手を使って技術を学

[162] Boyd et al., 2023, からの例。Automated Emotion Recognition in the Workplace: How Proposed Technologies Reveal Potential Futures of Work..

[163] 混合学習とは、デジタル（オンライン学習を含む）および非デジタル学習ツールの混合など、教育および訓練プロセスにおいて、複数のアプローチを取ることと理解される。

び[164]、かつ継続的な訓練を行う学校を含む[165]。教育機関は、通常、関係する国の教育当局または同等の当局によって、認定または認可される。主な特徴は、教育機関が証明書を発行できることである（それぞれ、加入は、証明書を得るための前提条件である）。この禁止事項は、入学許可プロセス中の候補者にも適用されると理解されなければならない。

たとえば：

-教育機関外において、言語学習のため、感情認識を用いたAIベースのアプリケーションをオンラインで使用することは、AI法第5条第1項(f)において禁止されない。これに対し、学生が教育機関からアプリケーションの使用を義務づけられている場合、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、オンラインで学生の試験を実施する際、AIベースの視線追跡ソフトウェアを使用し、視線の固定点と動き（たとえば、不許可の資料が使用されていないか探知するための注視点）を追跡することは、システムが感情を識別しまたは推測しないため、禁止されない。これに対し、システムが感情的な興奮や不安などの感情を探知するためにも使用される場合、これは禁止の範囲に含まれる。

- 教育機関が学生の興味や注意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、禁止される。これに対し、ロールプレイの文脈において学習目的でのみ導入される場合（たとえば、行為者や教師の訓練）、その結果が訓練を受ける人の評価や資格付与に影響を与えない場合、感情認識システムは許可される。

- 新入生の入学試験において、教育機関が感情認識AIシステムを使用することは禁止される。

- 教育機関がオンラインでの講義中、学生がその電話や他のチャンネルを通じてお互いに会話していることをとらえることができるAIシステムを使用することは、感情を推測しないため、禁止されない。これに対し、システムが、感情的な興奮、不安および興味などの感情を探知するためにも使用される場合、これは禁止の範囲に含まれる。

- 教育機関が、教師（職場）および学生（教育）の双方に感情認識AIシステムを用いることは禁止される。

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条第1項(f)の禁止事項は、治療用の使用のためのシステムのような、医療上または安全上の理由により職場および教育機関の分野において使用される感情認識システムに対する明

[164] たとえば、欧州委員会の提案に伴う影響評価は、職業訓練機関による特定のAIの使用が広範な基本的権利に対して強い干渉をもたらすと述べる。たとえば、欧州委員会、Commission Staff Working Document. Impact Assessment. Annexes, SWD(2021)84 final, Part2/2, p. 43. また、I. Tuomi, The , The use of Artificial Intelligence (AI) in education, European Parliament, 2020, pp. 9-10 も参照。

[165] 憲章第14条参照。

確な例外を含む。[166] ハイレベルな基本的権利の保護を確保するという AI 法の目的に照らして、この例外は狭く解釈されなければならない。

(257) 特に、治療用の使用は、CE マークが付された医療機器の使用を意味すると理解されなければならない。さらに、この例外は、健全性の一般的な側面を探知するための感情認識システムの使用を含まない。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、健康または安全の観点から許可されない。たとえば、職場または教育機関において燃え尽き症候群またはうつ病を探知することを目的とする AI システムは、例外の対象ではなく、依然として禁止される。

(258) この例外における安全上の理由の概念は、生命および健康の保護に関してのみ適用され、他の利益、たとえば盗難や詐欺から財産を保護するものではないと理解されなければならない。

(259) 例外のこの狭い解釈からすれば、医療上および安全上の理由によるあらゆる使用は、時間的制限、個人的な適用および規模を含む、厳密に必要かつ相応なものに依然として限定され、かつ十分な保護措置が伴うものでなければならない。このような保護措置には、たとえば、特定のユースケースに関する事前の書面による理由が付された専門家の意見を含み得る。必要性は、医療上および安全上の目的との関係で客観的に評価されなければならず、雇用主または教育機関の「ニーズ」を参照してはならない。この評価には、同じ目的を達する、より介入的でない代替手段が存在するかどうかを検討しなければならない。

(260) 雇用者および教育者は、明確な必要性がある場合に限り、医療上および安全上の理由により感情認識システムを導入しなければならない[167]。この文脈において収集されおよび処理されたデータは、他のいかなる目的にも使用することはできない。職場における AI 管理ソフトウェアの使用は労働者の健康および安全に潜在的に悪影響を与えることが証明されていることから、これは特に重要である。たとえば、ウェアラブル機器による継続的なモニタリングは、生産性に影響を与え、仕事上のストレスを増大させる可能性がある[168]。

(261) AI 法前文 18 項は、痛みや疲労のような身体的状態を感情認識システムの定義から除外しているため、安全上の理由により使用される多くの AI システム（たとえば、事故を防止する目的で職業パイロットまたはドライバーの疲労状態を探知するために使用されるシステムなど）は、もはやその定義に該当しない。

(262) データ保護ルールを含む他の法は、AI 法第 5 条第 1 項(f)における例外の要件を満たす感情認識システムに引き続き適用される[169]。

---

[166] AI 法前文 44 項。
[167] EU 労働法に従い、そのような新たな技術が導入された場合、雇用主は労働者またはその代表者らとも協議し、国内手続きに従わなければならない。これらの手続的要件を遵守することなく、AI 法自体を参考とすることによって、そのようなシステムを導入ことはできない。 また、データ保護法の観点からも同意が必要であり、これは依然として適用される。
[168] AI 法および EU の労働安全衛生上の枠組みとの相互関係 - Global Workplace Law & Policy(kluwerlawonline.com)
[169] 2026 年 12 月より、プラットフォーム業務における労働条件の改善に関する 2024 年 10 月 23 日の欧州議会および欧州理事会指令(EU)2024/2831 が適用される。

(263) AI 法第 6 条第 2 項および付属書 III、1(c)に従い、ハイ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 節のハイリスク要件および AI 法第 50 条第 3 項の透明性義務を遵守する必要がある。

たとえば： 感情認識は、自閉症の被用者や学生を支援したり、視聴覚障害者のアクセシビリティを向上させるために、医学上の理由により導入されることがある[170]。このような使用は、AI 法第 5 条第 1 項(f)の医療上の理由による例外に該当する。 これに対し、学生または被用者のウェルビーイング、モチベーションレベル、および仕事または学習の満足度を評価するための感情認識は、「医療上の理由による使用」に該当せず、禁止される。 雇用者は、たとえば危険な機械を導入したり、危険な化学物質を扱ったりする場合など、ストレスレベルの上昇／集中力の欠如が特定の危険をもたらす場合を除き、測定されたストレスレベルに基づいて不安を測定したり、被用者の倦怠感を測定するために、職場に AI 対応の機器またはデジタルアシスタントを導入することが禁止される。後者の場合、雇用者は、従業員の業績評価など、他の目的でデータを使用することはできない。

7.3. より有利な加盟国の法

(264) AI 法第 2 条第 11 項は、EU または加盟国が「雇用者による AI システムの使用に関する労働者の権利の保護に関し、労働者により有利な法律上、規則上、または行政上の規定」を維持しまたは導入することができることを規定する。労働者にとってより有利な団体交渉の協定も認められまたは奨励されることがある。

たとえば、加盟国は、労働分野における感情認識システムの使用を、医療上の目的で適用してはならないと規定する法を採択することができる。

7.4. 適用範囲外となるもの

(265) 上述のとおり、適用範囲外となるものは、次のとおりである：
- 生体データに基づかない感情および情緒を推測する AI システム、
- 痛みおよび疲労などの身体的状態を推測する AI システム。

(266) 職場および教育機関以外のあらゆる領域において使用される感情認識システムは、AI 法第 5 条第 1 項(f)の禁止事項に該当しない。しかし、このようなシステムは、ハイリスク AI システムとみなされる。[171] 同時に、そのようなシステムは、AI 法第 5 条第 1 項(a)および(b)（有害な操作および悪用）により、または他の EU 法により、特定の場合において禁止されることがある。EU

[170] システムは、被用者または学生/生徒が同僚などの感情を理解することを助けるために有用に用いられ得る。
[171] AI 法第 6 条第 2 項および附属書 III、1 (c)。

データ保護法、消費者保護など、その他あらゆる適用可能な法は、このようなシステムに対し引き続き適用される。

たとえば： 顧客対応のため商業上の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どうかにかかわらず、AI 法第 5 条第 1 項(f)の禁止に該当しない。したがって、たとえばキーストロークに基づき、または顧客の音声メッセージ（たとえば、チャットメッセージ、ヴァーチャル音声アシスタントの使用）に基づき、感情を認識できる AI システムなど、パーソナライズされたメッセージを表示するアプリケーションのためや、スマート環境（「インテリジェントビルボード」）を含む広告目的のために、オンラインマーケティングにおいて使用される例は、禁止事項に含まれない。 それにもかかわらず、そのような行為は、AI 法第 5 条第 1 項(a)および(b)の有害な操作および悪用の禁止の適用のためのすべての要件が満たされた場合、その禁止の対象となり得る[172]。

a）その他の適用範囲外のシステム

(267) 「群衆コントロール」とは、一般に、(公共の) 秩序およびイベントの安全性を維持するために、集団の行動を統制しおよびモニタリングすることをいう。それは大勢の人が集まるイベント（サッカーまたはフットボールの試合、コンサートなど）または、空港もしくは電車など、特定の場所に関係することが多い。群衆コントロールシステムは、たとえば、特定の場所での騒音や雰囲気のレベルを分析する場合に、個別の人の感情を推測することなく動作ができる。その場合、システムは（具体的な）自然人の感情を推測しないことから、AI 法第 5 条第 1 項(f)の適用範囲に該当しない。

(268) しかし、そのような群衆コントロールシステムは、たとえば、怒った顔が多いかどうかなど、個人の感情を推測することがあり得る。通常は、このような AI システムは、職場または教育機関においては使用されないことから、AI 法第 5 条第 1 項(f)の禁止事項には該当しない。

(269) また、たとえば、介護ロボットや、医療従事者がその職場において用いる感情認識システム、および緊急通報を分析する音声モニターなど、医療分野において使用されるシステムも適用範囲に入らない。

(270) このようなシステムは、たとえば、サッカースタジアムや中央駅（そのようなシステムが攻撃的な行動を認識するために使用される場所）の警備員、または医療分野における従業員など、多くの場合、仕事の状況においてそこにいる人をスクリーニングする。このような場合、導入者は、従業員のスクリーニングを回避するための保護措置を導入しなければならない。しかし、そのようなシステムがそれらの従業員の感情をも推論することは、完全に避けることはできない。このシステムの主な目的は従業員の感情の評価ではないため、これらのシステムは禁止の適用範囲外

[172] このような状況は、データ保護や消費者保護など、他のルールのもとでも禁止される場合がある。

であると考えなければならない。このようなシステムの導入者は、依然として、従業員がその使用により悪影響を受けないよう確保する責任がある。

8. AI 法第 5 条第 1 項(g):一定の「機微な」特性に対する生体分類

(271) AI 法第 5 条第 1 項(g)は、その人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向に関して推測または推論する目的で、その生体データに基づいて自然人を個別的に分類する生体分類システムを禁止する。この禁止事項は、たとえば法執行目的で使用され得る EU 法または国内法に従って取得された生体データセットのラベリング、フィルタリングまたは分類を対象とするものではない。[173]

8.1. 理論的根拠および目的

(272) 「機微」情報を含む広範な種類の情報は、それらの人を分類するために、関係者が知ることもなく、生体情報から抽出され、推定され、または推論される。これは、たとえば、誰かが一定の人種であるとみなされることによりサービスが拒否される場合のように、不公平かつ差別的な取扱いにつながり得る。性的指向もしくは政治的指向または人種などの側面に関連し、自然人を特定のグループまたはカテゴリーに割り当てることを目的とする AI ベースの生体分類システムは、人間の尊厳を侵害し、プライバシーや差別を受けないことなど、他の基本的権利に対し重大なリスクをもたらす。したがって、それらは AI 法第 5 条第 1 項(g)により禁止される。

8.2. 禁止事項の主な概念および構成要素

> AI 法第 5 条第 1 項(g)は、次のように規定する。
> AI に関する以下の行為は、禁止される:
> その人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向に関して推測または推論する目的で、その生体データに基づいて自然人を個別的に分類する生体分類システムを上市し、この特定の目的のためにサービスを開始し、または使用すること;この禁止は、法執行の分野における、生体データまたは生体データの分類に基づいて、画像のように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含まない。

(273) AI 法第 5 条第 1 項(g)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない:

(i)当該行為が、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」を構成すること;

(ii)当該システムが生体分類システムであること;

(iii)個々の人が分類されるものであること;

(iv)その生体データに基づくこと;

[173] AI 法前文 30 項。

(v) その人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活、または性的指向を推測または推論するためであること。

(274) 禁止事項が適用されるためには、５つの要件すべてが同時に満たされなければならない。最初の要件、すなわちAIシステムの上市、サービス開始または使用は、2.3において分析している。したがって、この禁止事項は、そのようなAIシステムを上市し、サービスを開始しまたは使用しないよう、AIシステムの提供者および導入者の双方に、それぞれがそれぞれの責任の範囲内において適用される。禁止事項の適用に関するその他の要件は[174] 、以下、さらに説明し、かつ分析する。

(275) この禁止事項は、法執行目的を含む、合法的に取得し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含まない。

8.2.1. 生体分類システム

(276) 「生体システムによる個人の分類は、通常、個人の生体データが事前に定義された特性を持つグループに所属しているかどうかを確立するプロセスである。個人を識別したり、その同一性を確認したりするのではなく、個人を一定のカテゴリーに割り当てることである。たとえば、広告ディスプレイは、その年齢や性別に基づき、それを見る個人により異なる広告を表示することがある。」[175] また、人は、統計的な理由により、識別されることなく、また、それらを識別する目的もなく、単に分類されることもある。

(277) AI法第3条(40)は、生体分類システムを、自然人の生体データに基づき、その自然人を一定の分類に割り当てることを目的とするAIシステム。それが他の商業サービスに付随するもので、かつ客観的な技術的理由のために厳密に必要な場合を除く、と定義する。7.2.1.d)における説明のように、「生体データ」は、AI法第3条(34)で定義される。特に、生体データは、生体の特徴に基づく行動上の特徴を含む。スカーフや十字架などの衣服やアクセサリー、およびソーシャルメディアでの活動による分類は、生体分類の範囲から除かれる。

(278) 生体分類は、それに基づき人が特定のカテゴリーに割り当てられる、身体的特徴（たとえば、顔の特徴や形、肌の色）のカテゴリーに依拠する場合がある。これらのカテゴリーの中には、特別の「機微な」性質のもの、または人種など、EUの反差別法のもとで保護される特性のものもある。しかし、生体分類は、DNAや、キーストローク分析または人の歩き方など、行動上の側面にも基づく場合がある[176]。

---

[174] 「AIシステム」、「上市」、「この特定の目的のためのサービス開始」、または使用の基準については、上記参照。
[175] Article 29 Working Party, Opinion 3/2012 on developments in biometric technologies, WP193, 27.4.2012, p6 参照。
[176] たとえば、Article 29 Working Party, Opinion 3/2012 on developments in biometric technologies, WP193, 27.4.2012, pp.16-17 参照。このグループは、ここでは「ソフト認識」(17頁)、すなわち「人々の行動または特定のニーズの検出」と述べる。

(279) AI 法に基づく生体分類の定義の範囲外となるには、2 つの要件 - 「他の商業サービスに付随するもので、かつ客観的な技術的理由のために厳密に必要であること」 - を累積的に満たさなければならない。

(280) AI 法前文 16 項によれば、純粋に付随的な特性とは、他の商用サービスに内在的に結びつけられる特性であり、その特性は客観的に技術的理由により主たるサービスなくして使用されず、および、この特性または機能の統合は AI 法のルールの適用可能性を回避する手段ではないことを意味する。

たとえば、AI 法第 5 条第 1 項(g)に基づき、次の AI の使用が認められる：

- 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オンラインマーケットプレイスで使用される顔や体の特性を分類するフィルターは、製品を販売することからなる主たるサービスに関連してのみ使用され得るものであるから、そのような付随的な特性を構成するものといえる。

- オンラインソーシャルネットワークサービスに統合されたフィルターで、ユーザーが写真またはビデオを追加しまたは変更できるように顔または身体の特性を分類するものも、それらは、オンラインでのコンテンツの共有からなるソーシャルネットワークサービスの主たるサービスなく使用され得ないものであるから、付随的な機能とみなすことができる。

これに対し、禁止される使用例は、次のものを含む：

- ソーシャルメディアプラットフォーム上で活動する人物を、その者がプラットフォームにアップロードした写真から生体データを分析することにより、その想定される政治的指向に従って分類し、それらに政治的ターゲティングメッセージを送信する AI システム。このようなシステムは、政治広告に付随するものに過ぎないかもしれないが、「客観的に技術的理由により厳密に必要」とはいえず、生体分類の定義から除外する条件は満たさない。

- ソーシャルメディアプラットフォーム上で活動する人物を、そのプラットフォームで共有された写真から生体データを分析することにより、その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者に広告を送信する AI システムは、AI 法の意味における生体分類に該当する。また、この場合も、この「付随的サービス」に厳密な必要性はないため、禁止事項からの除外は適用されない。

8.2.2. 人がその生体データに基づき個別的に分類されること

(281) 自然人の分類のための生体データの使用は、禁止事項が適用されるための不可欠の要素である（上記 8.2.1 および 7.2.1.d 参照）。

(282) さらに、禁止事項が適用されるためには、自然人は「個別的に」分類されなければならない。これが生体分類の目的または結果でない場合、たとえば、個人を見ることなくグループ全体を分類する場合、禁止事項は適用されない。

個別の分類の例には、次のものを含む： - 「属性推定」（人口統計を算出）する AI システム。顔、身長または肌、目および髪の色（またはそれらの組み合わせ）などの身体的特性に基づき、たとえば「年齢、性別、民族性」を含む。 - 個人を分類し、特定の特性（たとえば、右目の下の傷跡）に基づき、または右手に入れ墨があることにより、それらを選別できる AI システム。 これらのユースケースは、個々の生体分類の例である。これらの例が AI 法第 5 条第 1 項(g)の禁止に該当するためには、当該規定のすべての要件を満たさなければならない。

8.2.3. その人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向を推測または推論すること

(283) AI 法第 5 条第 1 項(g)は、限定された機微な特性を推測または推論することを目的とする生体分類システムのみを禁止している：つまり、人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向である。

たとえば、AI 法第 5 条第 1 項(g) に基づき禁止されるシステムは、次のものを含む： - その声から個人の人種を推測できる生体分類システム（これは、肌や目の色に従って人を分類するシステムや、犯罪被害者の出自を考慮してその DNA を分析するシステムとは異なる。これらのシステムは禁止されない）。 - その入れ墨や顔から個人の宗教的指向を推測できる生体分類システム。

8.3. 適用範囲外となるもの

(284) AI 法第 5 条第 1 項(g)の禁止事項は、法執行の分野を含む、生体データに基づく画像など、合法的に取得し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行う AI システムには適用されない。これは、AI 法前文 30 項においてさらに説明される[177]。

(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グループを等しく表し、たとえば、1 つの特定のグループを過度に表してい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得る。アルゴリズムのトレーニングに使用されるデータ

[177] AI 法前文 30 項：「この禁止は、たとえば法執行の分野において使用され得る髪の色や目の色による画像の選別など、生体データに基づき、EU 法または国内法に従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、または分類を対象とするものではない。」

が特定のグループに対して偏っている場合（すなわち、データの収集方法を原因としてデータにおける体系的な違いがグループ間に存在する場合、またはデータが史実上偏っている場合）、アルゴリズムはこのバイアスを再現する可能性があり、人や人のグループに対する違法な差別をもたらし得る。[178] このため、保護された機微情報に基づくラベリングは、高品質なデータのため、さらにいえば、差別を防ぐために、必要になる場合がある。AI 法は、ハイリスク AI システムに関する AI 法の要件遵守のため、ラベリング操作を要求することすらあり得る。[179]したがって、このような生体データのラベリングまたはフィルタリングは、AI 法第 5 条第 1 項(g)における禁止から明示的に除外される。この禁止事項は、人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向を推測するために生体データが分類される場合にのみ適用される。

認められ得るラベリングまたはフィルタリングの例は、次のものを含む：

- 特定のグループの実績が悪い、つまり他のグループよりも結果が悪いデータに基づきアルゴリズムが「訓練」されたため、ある民族グループのメンバーが就職面接に招かれる可能性が低くなるケースを避けるための生体データのラベリング。[180]

- 皮膚や目の色により画像を使用して患者を分類することは、がんの診断など、医療診断にとって重要となり得る。

(286) AI 法第 5 条第 1 項(f)はまた、当該規定における禁止事項が、法執行分野において合法的に取得したデータセットのラベリングまたはフィルタリングには適用されないと規定する[181]。

たとえば、これは、子どもの性的虐待素材が含まれているとの疑いがあるデータセットのラベリングおよびフィルタリングを可能にする AI システムの、法執行機関による使用を対象とする。最初の段階においては、法執行機関は、AI システムの支援により、画像から機微データを検出し、編集する。さらに、性別、年齢、目や髪の色、傷跡およびマーキングなどの生体データによるフィルタリングとラベリングは、被害者の特定や他の事件との関連づけに役立ち得る。同様に、指の長さまたは識別用のマーキングや入れ墨などの具体的特性に基づき、加害者の手をフィルタリングおよびラベリングすることは、被疑者特定のために許可される。

8.4. 他の EU 法との相互作用

(287) 生体データを基礎として GDPR 第 9 条第 1 項により保護される機微な属性または特性に従い生体分類のための使用が意図される AI システムは、本規則に基づき禁止されない限り、[182]AI 法[183]に基づきハイリスクに分類される。

[178] 同上。
[179] たとえば、AI 法第 10 条および第 17 条参照。
[180] FRA、# BigData. Discrimination in data supported decision making, Luxemburg, 2018, 14, p. 5.
[181] 法執行のための生体分類システムの使用に関する AI 法は、TFEU 第 16 条に基づく。AI 法前文 3 項も参照。
[182] AI 法前文 54 項および附属書 III、1、b)。
[183] AI 法前文 54 項および附属書 III、1、b)。

(288) AI 法第 5 条第 1 項(g)は、GDPR、LED、EUDPR などの EU データ保護法に基づく合法的な個人データ処理の可能性をさらに制限する。特に、AI 法第 5 条第 1 項(g)は、上述のとおり、法執行の分野を含む、合法的に取得し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別として、人種、政治的意見、労働組合への加入、宗教上または思想上の信念、性生活または性的指向を推測するための、AI 法で定義されている生体認証データに基づく、自然人の生体分類の可能性を除外する。さらに、AI 法第 5 条第 1 項(g)の禁止事項は、人種、民族的出自、性的指向、政治的意見、宗教上の信念など、特別なカテゴリーの個人データに基づく差別をもたらす「プロファイリング」を明示的に禁止する LED 第 11 条第 3 項と整合する。

9. AI 法第 5 条第 1 項(h) - 法執行目的のためのリアルタイム遠隔生体識別(RBI)システム

(289) AI 法第 5 条第 1 項(h)は、AI 法において限定列挙されている例外を除き、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI システムの使用を禁止する。具体的には、AI 法第 5 条第 1 項(h)(i)ないし(iii)は 3 つの状況を想定し、国内法により許可され、かつ AI 法第 5 条第 2 項ないし第 7 項の要件および保護措置を満たす場合に、そのようなシステムの使用が許可される。

(290) AI 法第 5 条第 5 項に従い、加盟国は、その領土内において、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI システムの使用が許可されるかどうか、また 3 つの状況のうちどの状況で許可されるかを自由に決定できる。そのような使用を許可しおよび規制する国内法がない場合、法執行機関およびその代理として行動する主体は、法の執行を目的としてそのようなシステムを導入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の存在は、そのような使用の前提条件である。

(291) AI 法第 5 条第 1 項(h)は、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI システムの使用のみを禁止しているため、その規定に関係するのは、そのようなシステムの導入者のみである。そのようなシステムの上市およびサービス開始、ならびにその他の RBI システムの使用は禁止されていないが、AI 法第 6 条第 2 項および付属書 III (a)に従い、ハイリスク AI システムのルールの適用対象となる[184]。加盟国が、AI 法第 5 条第 1 項(h)に列挙される 3 つの目的のいずれかのために、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI システムの使用を許可した場合、ハイリスク AI システムのルールもその使用に適用される。

(292) 最後に、特別なルールが、法の執行を目的とする RBI システムの遡及的使用に適用される。このようなリアルタイムでない使用は禁止されないが、ハイリスク AI システムの導入については、追加的な保護措置が適用される（AI 法第 26 条第 10）。

9.1.理論的根拠および目的

[184] さらに、法の執行を目的とする RBI システムの遡及的使用に適用される具体的ルール（AI 法第 26 条第 10 項）。

(293) AI 法前文 32 項は、次のようにいう。「法の執行を目的とする、公衆がアクセス可能な場所内における自然人の「リアルタイム」RBI のための AI システムの使用は、人口の大部分のプライバシーに影響を与え、絶え間ない監視の感覚を想起させ、集会の自由およびその他の基本的権利の行使を間接的に思いとどまらせる可能性があることにより、関係者の権利および自由に特に介入する。自然人の遠隔生体識別を目的とする AI システムの技術的不正確性は、バイアスがかかった結果をもたらし、差別的な影響を引き起こす可能性がある。このバイアスがかかった結果や差別的な影響のリスクは、年齢、民族的属性、人種、性別、心身障害に関して、特に重大である。さらに、リアルタイムで運用されるシステムの使用は、影響が即時である事実および追加的検証または訂正を行う可能性が限られる事実により、法執行行為の文脈における関係者、またはそれによって影響を受ける関係者の権利および自由に対するリスクを増大させる。」

(294) しかし、そのようなシステムの使用が、実質的な一般的利益を達成するために必要最小限であり、かつそのような使用が起こり得る状況を限定列挙し、厳格に定義している場合においては、当該使用は、基本的権利に対するリスクに勝る（AI 法前文 33 項)。このようなシステムが「責任ある適切な方法」において使用されることを確保するため、それらの使用は、AI 法第 5 条第 2 項ないし第 7 項のセーフガードならびに特定の義務および要件の対象となる。

9.2. 禁止事項の主な概念および構成要素

AI 法第 5 条第 1 項(h)

AI に関する以下の行為は、禁止される：

(h) 法の執行を目的として、公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用すること。 以下の目的のいずれかに鑑み、この使用が必要最小限であり、かつ必要最小限の範囲内である場合を除く：

(i) 誘拐、人身売買または性的搾取の特定の被害者を対象とする捜索、および行方不明者の捜索；

(ii) 自然人の生命もしくは身体の安全に対する特定の、具体的かつ差し迫った脅威の防止、またはテロリスト攻撃の現実かつ現在の、もしくは現実かつ予見可能な脅威の防止；

(iii) 附属書Ⅱに定める犯罪であり、かつ、関係する加盟国において最長期間が少なくとも 4 年間である拘禁刑または留置命令によって処罰される犯罪に対する、刑事捜査、訴追または刑事罰の執行を目的とする、刑事犯罪を犯した被疑者の所在の特定または身元の特定。

第 1 項(h)は、法の執行以外の目的による生体データ処理のための規則(EU) 2016/679 第 9 条を害しない。

(295) AI 法第 5 条第 1 項(h)の禁止事項が適用されるためには、いくつかの累積的要件を満たさなければならない：

(i) AI システムは RBI システムであること；

(ii) その活動が、そのシステムの「使用」により成ること；

(iii)「リアルタイム」であること；
(iv) 公衆がアクセス可能なスペースであること、および
(v) 法執行目的であること。

(296) 第2の要件、すなわちAIシステムの「使用」は、既に本ガイドライン2.3で分析した。上記に列挙する他の要件は、以下でさらに説明しかつ分析する。

9.2.1. 遠隔生体識別の概念

(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）または行動上の特徴（歩き方や声など）を探知し、取得し、および機械により読み取り可能な生体データ（上記7.2.1.d参照）に変換する。これらのデータは、さまざまな形式で利用可能である：認識目的で使用される、画像や個人の顕著な特徴を数学的に表現したテンプレートなど。生体認識技術は、確認および識別の目的で使用される。[185]

(298) AI法第3条（41）によれば、RBIシステムは、次のとおりである。

自然人が能動的に関与することなく、個人の生体認証データと参照データベース内にある生体データとを比較することにより、通常は遠隔で、自然人を識別することを目的とするAIシステム

(299) この定義は、生体認識システムの識別機能のみを対象とし、これは関係者の能動的な関与がないこと（すなわち、積極的な参加がないこと）を意味し、かつ通常は、離れた場所からその者の特性を取得する結果になる。識別実行のため、取得された生体データは、参照データベース（たとえば、顔画像や被疑者のテンプレートを含む犯罪データベースなどのリポジトリ）に既に保存されている生体データと比較される。

a) 識別目的のみ

(300) 「生体識別」の概念は、AI法第3条(35)において、次のように定義される。

自然人の生体データをデータベース内に保存されているその生体データと比較することにより、自然人の識別を行う目的で、人の身体的、生理的、行動的、または心理的特徴を自動認識すること

(301) AI法前文15項は、そのような人間の特徴が以下を含み得ることを明らかにする。

顔、眼球運動、体型、声、抑揚、歩き方、姿勢、心拍数、血圧、匂い、キーボードのタッチ

(302) 自然人を追跡するために使用されるAIシステムもまた、生体識別の定義に含めることができる。たとえば、被疑者が逃走する方向を確認することである。これは、犯罪の被疑者の所在特定

[185] ISO/IEC 規格 2382-37 における生体コミュニティによる定義 :2022 情報技術–ボキャブラリー、生体認識、用語 37.01.03

を認める AI 法第 5 条第 1 項(h)(iii)による結論である。所在特定は、人が追跡されている場合に可能となる。

(303) 生体確認のための使用を意図する AI システムは、AI 法第 5 条第 1 項(h)の禁止事項の範囲外となる。[186] 生体確認（または認証）は、センサーで示されたデータを、スマートフォン、パスポート、または ID カードなどのデバイスに保存された事前に記録された別のデータセットと比較することからなる。生体確認の目的は、特定の人物がそれと主張する本人であることを確認することである。

生体確認の例は、スキャンされた旅行者の顔とそのパスポートに含まれる顔画像との e ゲートでの比較である。

b）遠隔性

(304) AI 法第 3 条(41)によれば、遠隔性とは、個人が能動的に関与することなく、生体システムが、通常は、遠隔で、個人の生体データと参照データベース内にある生体データを比較することにより、個人を識別できる能力を意味する。

(305) サービスへのアクセス、デバイスのロック解除、または施設へのセキュリティアクセスを得ることのみを目的とする、自然人の同一性を確認するための生体システムの使用は、「リモート」の概念から除外される（AI 法前文 15 項）。この方法は、たとえばアクセスコントロールにおいて用いられる。[187]

たとえば、顔識別システムは、顔スキャン技術を通じ、制限区域（たとえば発電所の敷地）に立ち入るために導入される。このシステムは、エントランスカメラに表示された個人の顔と、建物への入館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を比較照合する。

(306) AI 法前文 17 項が明確にするところによれば、禁止の適用範囲からのこの除外は、そのようなシステムが、自然人の積極的な関与なく多数の人の生体データの処理に使用され得る RBI システムと比較し、自然人の基本的権利に対し与える影響が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。さらに、当該前文が明確にするところによれば、RBI システムは、通常、複数の人またはその行動を同時に認識するために使用され、その能動的な関与なく、自然人の識別を大幅に容易にするために使用される。能動的な関与のためには、人がカメラの存在について知らされるだけでは不十分であり、能動的な関与を促進する方法で設置されたカメラの前で、積極的かつ意識的に歩み寄る必要がある。

たとえば - 監視目的で地下鉄駅の壁や天井に設置されたカメラに使用される RBI システム。このようなシステムは、遠隔性の条件を満たす。

[186] AI 法前文 17 項。
[187] 例として、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.

- 地下鉄の駅へのアクセスに使用されるシステム。たとえば地下鉄のバイオメトリック・チケットなど、人が能動的に関与し、意識的に生体センサーに近づいてアクセスを得るシステムは、その条件を満たさない。

(307) 指紋、歩き方、声、DNA、キーボードのタッチ、その他の（生体上の）行動信号を処理する生体認識システムもまた、RBI システムを構成し得る。[188]

たとえば、
- 声紋の生体技術システムは、話者を識別するために導入され得る。そこで、マイクは生体サンプルを収集する。
- CCTV を通じ歩き方の認識システムが使用され得る。そして、ビデオは従前キャプチャされたテンプレートと一致するかどうか自動的にチェックされる。
- キーボードのタッチの生体技術は、不正なメッセージを入力している者を特定するために使用され得る。

これらのシステムが RBI システムの例に挙げられるという事実は、それらが AI 法第 5 条のもとで禁止されることを意味しない。

(308) 個別の法執行者により使用される RBI を可能とするボディカメラの場合、たとえば、数百人の参加者が参加するデモにおける無差別の撮影は、遠隔性の条件を満たすとみなされる。

c) 参照データベース

(309) 比較のための生体データを含む参照データベースがなければ、識別は不可能である。したがって、識別目的で比較を行うためには、参照データベースの存在は不可欠である。[189]

たとえば、行方不明者の場合、シェンゲン情報システムの[190] データベースは、（運用開始後）顔認識目的で参照データベースとして使用され得る。

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、「まったくの瞬時、ほぼ瞬時に、またはいかなる場合にも大きな時間的格差なく」[191]取得し、さらに処理することを意味する。 すべての処理のステップ、すなわち、生体データの取得、比較、および識別は、同時にまたはほぼ同時に行われ、これには、RBI システムの遡及的使用を通じ、禁止の迂回を回避するための「限定的な短い時間的格差」が含まれ得る。[192] 「大きな時間的格差なく」との概念は、AI 法において定義されていない；これはケースバイケースで評価されなければならない。リアルタイムまたはリモート後の識別に用いられるデバイスは、ますます機能が異なる同じものとなっているため、区別は一時的

[188] EDPB-EDPS、共同意見 5/2021、11 頁。欧州理事会、「Opinion of the Legal Service」12302/22、2022 年 9 月 12 日、33 項、および AI 法前文 15 項。
[189] AI 法前文 34 項。
[190] 行方不明者に関するアラート（SIS II 決定第 32 条）；第 2 世代シェンゲン情報システム(SIS II)の創設、運用および使用に関する 2007 年 6 月 12 日の欧州理事会決定 2007/533/JHA
[191] AI 法前文 17 項。
[192] AI 法第 3 条（42）

なものである。一般的に言えば、時間的格差は、少なくとも、生体データが取得された場所をその人が去った可能性がある場合に重要である。

(311) 一般に、リアルタイムシステムは、特定の場所で、人を遡及的に識別するためではなく、迅速な対応を容易にするために用いられる。これらは、システムのユーザーに対し、監視下にある人の動きを追跡し、かつ監視する手段を提供する。

a)コンサート会場へのすべての来場者をスクリーニングするAIシステム:リアルタイムRBI
b) コンサートへのすべての来場者を撮影するシステム。コンサートにおいて事件が起こり、コンサート後、犯罪者を特定するために、ビデオ素材に対して識別システムを運用する場合：事後的RBI。

(312) 法執行機関がモバイルデバイスにより秘密裏に人の写真を撮影し、それを即時検索のためにデータベースに送る場合、状況によっては、AI法第5条第1項(h)の禁止に該当し得る。

9.2.3. 公衆がアクセス可能な場所

(313) AI法第3条(44)は、公衆がアクセス可能な場所を、次のように定義する。

適用され得る当該場所への何らかのアクセス条件が存在するかどうか、また潜在的な収容人数の制限にかかわらず、不特定数の自然人がアクセス可能な、公有または私有のあらゆる物理的場所。

(314) AI法前文19項は、そのような場所の特徴であるいくつかの要素を列挙する：

- たとえば、チケット購入や交通手段の購入、事前登録または一定の年齢であることなど、収容人数やセキュリティ制限を問わない、不特定数の人のアクセス可能性。施錠されていないドアからある場所へのアクセス可能性は、表示や状況が反対の事実（たとえばアクセス制限の標識など）を示唆しているならば、その場所は公衆によりアクセス可能であることを意味しない。さらに、その場所へのアクセスは、法律により定義されるとおり、公共の安全もしくはセキュリティに関係するまたはその場所に関係ある権限を有する人の決定に関係する、一定の人に制限し得る。

たとえば、公衆がアクセス可能な場所は、原則として次のとおりである；
- 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て見本市が開催されるイベント会場。

門で閉じられた場所は、門が施錠されていない場合であっても、たとえば、フェンス囲いがある数軒の居住エリアの門がある入り口など、通常、公衆がアクセス可能な場所とはみなされない。これに対し、公衆への開放時間がある居住地門内の公園で、アクセス制限がない時間中は、通常、その時間帯は公衆がアクセス可能な場所を構成し、その時間外は閉鎖された場所を構成する。

- 所有権は無関係である。すなわち、公衆がアクセス可能な場所とみなされるために、場所が公有である必要はない。

たとえば、その場所は、民間の主体が所有する場合、公共の主体が所有する場合、または公共の主体が所有しかつ民間の主体が管理する場合があり、場所の性質に影響を与えない。

- そのために場所が使用される特定の活動はない；公衆がアクセス可能なエリアは、必ずしも公共サービスに関連する場所ではない。さらに、公共サービスに関連する場所は、公衆がアクセス可能でない場所、つまり自治体で働く公務員のオフィスが含まれ得る。

たとえば、公衆がアクセス可能な場所は、店舗、レストラン、カフェなど、商取引に使用され得る；たとえば、銀行、専門職の活動（診療所や会計事務所）、ホスピタリティ（たとえば、ホテル）など、サービスに使用され得る；たとえば、スイミングプール、ジム、スタジアムなど、スポーツに使用され得る；たとえば、バス、地下鉄の駅や鉄道の駅、空港、交通手段など、輸送に使用され得る；たとえば、映画館、劇場、博物館、コンサートホールや会議場など、娯楽に使用され得る；または、道路や広場、公園、森林、遊び場など、レジャーやその他のために使用され得る。[193]

(315) 以下の場所は、AI 法第 5 条第 1 項(h)の意味における公衆がアクセス可能な場所を構成しない：

- オンラインの空間。AI 法第 3 条(44)にいう物理的な場所を構成しないからである。

たとえば、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、禁止の範囲から除外される。

- 立ち入りが管理され、または関連する従業員やサービス提供者に制限される、工場、企業、職場などの一定の場所は、限られた人数の者によるアクセスを意図している。それらは、関係する従業員やサービス提供者のみのアクセスが意図されるからである[194]。

たとえば、バッジでアクセスできる職場は、原則として公衆がアクセス可能な場所とはみなされないが、立ち入りの管理がないオフィスは、これに該当し得る。

- 刑務所と出入国管理局は、公衆がアクセス可能な場所ではない。[195]

(316) たとえば、国境検問所は、公衆がアクセス可能な場所ではないが、国境検問所に通じる道またはその近くの森林は、通常、公衆がアクセス可能な場所にあたる。

(317) 一部の場所は、二重の機能を持ち得る。たとえば、空港は、一般的に、共用エリアとして、公衆がアクセス可能な場所とみなされるが、出入国管理専用のエリア（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 法前文 19 項に明記されるとおり、場所が公衆にとってアクセス可能かどうかの評価は、ケースバイケースの分析に基づいて行われなければならない。

---

[193] AI 法前文 19 項。
[194] AI 法前文 19 項。
[195] AI 法前文 19 項。別の文脈では、国境管理は、規則(EU)2016/399(シェンゲン協定国国境法)に従って、その目的のために、国境を越える意図または越える行為にのみ対応して、国境で実施される活動として定義されています。これは、国境の両側で最大 50 キロメートルまで広がる可能性のある、いわゆる国境地域を含んでいません。

9.2.4. 法の執行を目的とすること

(319) AI 法第 5 条第 1 項(h)の禁止事項は、法執行活動を行う主体、当局、または組織に関係なく、法の執行を目的とする RBI システムの使用に適用される。

(320) 法執行は、AI 法第 3 条（46）において、「刑事犯罪の防止、捜査、探知もしくは訴追、または刑事罰の執行を目的として、法執行機関によって行われ、または法執行機関に代わって行われる行為」と定義され、これには公共の安全に対する脅威からの保護および当該脅威の防止を含む。これらの目的は、LED 第 1 条に列挙されているものと同じである。[196] したがって、LED に関連するこれらの目的の解釈はいずれも、AI 法において使用される「法執行」の概念を解釈する目的にも関連し得る。

(321) 法執行目的は、刑事犯罪の捜査、探知、および訴追を含む。また、それらは、なんらかの犯罪が実際に行われる前の、公共の安全に対する脅威に対する保護および防止を含む、犯罪の防止に関連する活動も含む。たとえば、警察は、犯罪防止の文脈において、「デモ、主要なスポーツイベント、暴動時における強制的措置」を取ることができる。[197] 最後に、これらの活動は、刑の執行などの罰則の執行を含む。

(322) AI 法第 3 条（46）によれば、法執行活動は、法執行機関またはそれに代わる者が実行し得る。法執行機関は、さらに、LED における国内の管轄当局の定義と同じように、AI 法第 3 条(45)において定義される。[198] この定義は、法執行機関、および委託組織または主体（民間の当事者の可能性がある）を対象とする。

(a) 公共の安全に対する脅威からの保護および当該脅威の防止を含む、刑事犯罪の防止、捜査、探知もしくは訴追、または刑事罰の執行を行う管轄を有するあらゆる公的機関；または

たとえば、このような公的機関には、法執行業務を遂行する場合の、警察当局や刑事司法当局（検察官など）を含む。

(b) 公共の安全に対する脅威からの保護および当該脅威の防止を含む、刑事犯罪の防止および探知、捜査または訴追または刑事罰の執行を目的として、加盟国の法が、公的機関の権限行使および公権力を委ねるその他のあらゆる組織または主体；

(323) AI 法に基づき、他の主体、組織、または個人は、上記に列挙する目的のため、加盟国の法により、公的権限および公権力を委託された後、法執行活動を行うことができる。

---

[196] 管理業務（人事など）を行う場合など、法執行機関の一部の活動は、LED の適用範囲から除外され、これらの活動は法執行の枠組みの外で実行される。それらは GDPR の適用となる。GDPR 前文 19 項参照。
[197] LED 前文 12 項。
[198] LED 第 3 条(7)。

(324)　「代わって（On behalf of）」とは、法執行機関が、法執行活動（またはその一部）の実施を他の主体または人（民間の者を含む）に委任するか、または特定の場合において、他の主体または人に法執行活動を支援する行動を要請することを意味する。いずれの場合も、法執行機関は、すべての主要な側面について指示し、かつ他の主体を監督しなければならない。これは、この要件が人に「代わって」行動するという概念に内在するためである。

他の者への任務の委任は、たとえば、以下を含む。 -法執行機関の指示および監督の下、公共交通網のセキュリティを確保するよう法執行機関から要請された公共交通事業者。 -法執行機関の指示および監督の下、スポーツイベントでセキュリティを提供することを法執行機関から要請されたスポーツ連盟。 -法執行機関の指示および監督の下、「特定の事件における一定の犯罪に対抗する」ための一定の行動をとることを法執行機関から要請された銀行。 これらの者は法執行機関に「代わって」これらの活動を行うため、これらの活動は「法執行の目的」の定義に該当する。これらの主体が犯罪（詐欺、マネーロンダリングなど）を探知しおよび犯罪に対抗する場合に「自己のために」行動するのであれば、AI 法第 5 条第 1 項(h)の禁止に該当するとみなされない。

(325)　これらの他の組織または主体が特定の法執行業務を委任された場合にのみ、それらの活動は「法執行」の定義に該当する。

9.3. 禁止事項の例外

(326)　AI 法は、法の執行を目的とする公衆がアクセス可能な場所内でのリアルタイム RBI の使用に関する一般的な禁止に対して、3 つの例外を規定する。AI 法第 5 条第 1 項(h)(i)ないし(iii)は、リアルタイム RBI が許可され得る 3 つの目的を限定列挙し、AI 法第 5 条第 2 項ないし第 5 条第 7 項は、そのような許可の要件および保護措置を定める。AI 法第 5 条第 1 項(h)(i)ないし(iii)自体は、公衆がアクセス可能な場所内における RBI システムのリアルタイムの使用の法的根拠を構成するものではない。むしろ、特に、AI 法第 5 条第 2 項ないし第 7 項の要件を満たす加盟国の国内法のみが、AI 法第 5 条第 2 項に規定するとおり、リアルタイム RBI の使用を許可できる。
したがって、これらの目的の 1 つ以上に対するリアルタイム RBI の使用を許可する加盟国の法がない場合、2025 年 2 月 2 日以降、そのような使用は禁止される。

9.3.1. 理論的根拠および目的

(327)　AI 法第 5 条第 1 項(h)(i)ないし(iii)に定める目的は、法の執行を目的とする一定の AI および捜査ツールの使用を許可することを目的とする。これらの目的は、次のとおりである：

(i) 3 種の特定の重大犯罪の被害者および行方不明者を対象とした捜索[保護]；

(ii) 生命もしくは身体の安全に対する差し迫った脅威、またはテロリスト攻撃の現実の脅威の防止[予防]；および

(iii) 附属書 II に列挙される一定の重大犯罪の容疑者および犯罪者の所在の特定および身元の特定[訴追／捜査]。

(328) これらの状況において、EU 議会は、リアルタイム RBI システムがそれらのシステムの対象となる個人の基本的権利に生じるリスクに対し、社会の安全保障の必要性との均衡を図る。AI 法前文 33 項によれば、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI システムの使用が許可される目的は、厳格で、限定的かつ狭義に定められなければならず、かつ、基本的権利に生じる「リスクに勝る」「具体的な公共の利益」を達成するための「厳格な必要性」がある場合に現れる。AI 法第 5 条第 1 項(i)ないし(iii)に列挙されない、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI システムのその他のいかなる使用も禁止される。

たとえば、警察が万引き犯を特定するためリアルタイム RBI システムを使用し、かつそれらの顔画像を犯罪データベースと比較することは、AI 法第 5 条第 1 項(h)(i)ないし(iii)に列挙される目的のいずれにも該当しないため、禁止される。

9.3.2. 3 つの重大犯罪の被害者および行方不明者を対象とする捜索

(329) AI 法第 5 条第 1 項(h)(i)によれば、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI の使用は、厳密な必要性および AI 法第 5 条第 2 項ないし第 7 項の要件を前提として、誘拐、人身売買、または人の性的搾取の被害者を対象とする捜索、および行方不明者の捜索のために許可される。

a) 3 種の犯罪の被害者を対象とする捜索

(330) AI 法第 5 条第 1 項(h)(i)に定めるシナリオは、法執行機関が 3 つの重大犯罪の被害者捜索を支援することを求めている。

(331) 捜索対象には、被害者の所在の特定および身元の特定が含まれる。

3 種類の犯罪

(332) 3 つの重大犯罪の特定の被害者を対象とした捜索は、AI 法第 5 条第 1 項(h)(i)に列挙されるシナリオに含まれる：すなわち、誘拐、人身売買および性的搾取である[199]。

たとえば、子どもが誘拐され、誘拐犯が子どもをある場所から別の場所に車で連れ去ろうとしているという具体的な兆候がある場合、警察は、その子どもを捜索対象としてリアルタイム RBI システムを使用することがあるが、子どもを特定するためには、導入区域および使用期間を定めなければならない。

[199] 誘拐、人身売買、性的搾取は、犯罪被疑者または有罪判決を受けた者を逮捕し、かつ欧州逮捕状（EAW）を発行した国に移送する EAW の対象となり得る 3 つの犯罪である。3 つの犯罪は、主に女性と子どもに関連するが、これらに限定されない。欧州委員会の移民・内務総局によれば、被害者の約 40%は EU 市民であり、そのほとんどが性的搾取のために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、強制労働や強制的な物乞いのために人身売買されるが、ほとんどの女性や子どもは性的搾取のために人身売買される。https://home-affairs.ec.europa.eu/policies/internal-security/organised-crime-and-human-trafficking/together-against-traffickinghuman-beings_en

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対象とする。[200]

(334) 行方不明の子どもと行方不明のおとなとは、区別することができる。行方不明のおとなの自発的な失踪は、必ずしも捜索開始の要因になるとは限らないからである。行方不明の子どもに関して適用されるルールは、加盟国によって大きく異なる。[201] いずれにせよ、AI 法第 5 条第 1 項(h)(i)は、法の執行を目的として行方不明者を捜索するためのリアルタイム RBI システムの使用のみを許可する。

(335) おとなの行方不明は、大人には行方をくらます権利があるため、必ずしも警察によるその人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（「被後見」）、健康状態（精神疾患）、自殺メモの存在に関連するだけでなく、私物をもたずに出発したことにも関連する可能性がある。行方不明の状況が懸念される場合、捜索開始のために、警察に行方不明届けを提出し得る。

(336) 一部の加盟国においては、行方不明者の捜索は、法執行目的ではなく、行政手続きの下で行われる場合がある。たとえば、脆弱な者が行方不明となっている場合、犯罪または他の法執行目的の疑いがなければ、その者を捜索するためのリアルタイム RBI システムの使用は、法執行目的とはみなされず、GDPR に基づきそのような使用のためのルールが適用される。

9.3.3. 生命に対する差し迫った脅威またはテロリスト攻撃の防止

(337) AI 法第 5 条第 1 項(h)(ii)は、厳格な必要性および AI 法第 5 条第 2 項ないし第 7 項に定める要件を前提として、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI の使用が許可される第 2 のシナリオを掲げる。

自然人の生命もしくは身体の安全に対する特定の、具体的かつ差し迫った脅威の防止、またはテロリスト攻撃の現実かつ現在の、もしくは現実かつ予見可能な脅威の防止

(338)

a) 自然人の生命または身体の安全に対する特定の、具体的かつ差し迫った脅威

(339) 生命に対する権利を保障する憲章第 2 条の適用により、EU およびその加盟国は、個人の生命を保護しなければならない。公衆がアクセス可能な場所内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の生命への脅威に関する AI 法第 5 条第 1 項(h)(ii)の基準は、(1)特定の、(2)

[200] 「行方不明者」は EU レベルでは定義されていない。しかし、2021 年 12 月の「行方不明者の地域における国境を越えた警察協力の強化」に関する欧州理事会の結論において、欧州理事会は、欧州理事会勧告 CM/Rec(2009)12 および国内規則の双方における、行方不明者の定義を参考とする。欧州理事会の結論(2021)14808/21、11 項、4 頁。

[201] 欧州委員会、欧州移民ネットワーク、「EU 加盟国は付き添いのない未成年者の行方不明事件をどのように扱うか」EMN inform、2020 年。

具体的、かつ(3)生命または身体の安全に対する差し迫った脅威、(4)自然人の存在を要件とする。脅威は、特定された個人またはグループに限定される必要はなく、一般的に、自然人に関連する。

(340) AI 法前文 33 項は、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、重要なインフラへの差し迫った脅威も含まれ得ることを、次のとおり明確にする[202]。「この重要インフラの停止または破壊は、特に、住民への基本的物資の供給または国家の本質的機能の実行に対する重大な侵害によって、人の生命または身体の安全に対する差し迫った脅威をもたらし得る。」

たとえば、[203] 重要なインフラ（たとえば、発電所、水道、または病院）の重大な停止および破壊は、住民への基本的供給停止による深刻な害がある場合（特に、暖かいまたは寒い天候における、長期間にわたる電気または飲料水の喪失など）、人の生命や身体の安全に差し迫った脅威が生じ得る。

(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、特に、AI 法第 5 条の主要な要素および理論を考慮し、EU 法に従い、加盟国の国内法に基づいて最終的に定義されかつ評価される。これは、公衆がアクセス可能な場所内において法の執行を目的としてリアルタイム RBI を使用することの禁止に対する例外を用いるため、加盟国が採用しなければならない法において規定するか／言及する必要がある。

(342) 生命または身体の安全に対する**差し迫った**脅威とは、いつでも起こり得る脅威であり、「直ちに対応する」必要がある。[204] 身体の安全に対する**具体的**脅威は、重傷を負う場合に関係する。

(343) 特定の脅威とは、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し、それは、仮定的なものであったり、または一般的な一定の危険に関連するものであってはならない。

たとえば、元学生が数人の元クラスメートに復讐するために、かつての大学で致命的な攻撃を計画していることの情報を、警察が得た。警察は、襲撃の差し迫った状況、標的にされた学校、および計画を実行するために使用する予定の武器についての情報を受け取る。

(344) 特定の脅威は、故意である必要はない。故意でない行為は、生命または身体の安全を脅かす可能性がある。

b) テロリスト攻撃の現実かつ現在の、または現実かつ予見可能な脅威

(345) AI 法第 5 条第 1 項(h)(ii)に定める第 2 のシナリオのこの部分は、いくつかの要素からなる：テロリスト攻撃の脅威の存在および脅威の特性であり、これらは現実かつ現在の、または現実かつ予見可能でなければならない。

テロリスト攻撃の脅威

---

[202] 指令 2022/2557 第 2 条(4)に定義されるとおり
[203] AI 法前文第 33 項。
[204] 規則 2023/1543 の前文第 37 項。

(346)　脅威の存在および重大性に関する評価は、国家の安全保障、より具体的には、テロリスト攻撃の場合に、保障のため講じられるべき措置の実際の状況を評価する際、国家レベルで行われる。**テロリストの脅威のレベルは、国家レベルで定義され**、加盟国により異なる。たとえば、オランダ[205]は5つのレベルの脅威、ベルギー[206]は4つのレベル、フランス[207]は3つのレベル、スウェーデン[208]は5つのレベルを設定している。しかし、第5条第1項(h)(ii)で用いられる「現実かつ現在の、または現実かつ予見可能な脅威」の概念は、EU法の自律的概念であり、したがって、原則として、各国の定義から独立して評価されなければならない。脅威は、テロリズム全般に関連するものではなく、特に、テロリスト攻撃の脅威に関係する。

脅威の特性：現実かつ現在の、または現実かつ予見可能であること

(347)　法の執行を目的とする公衆がアクセス可能な場所内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達すべき重大性の閾値は、特に、テロリスト攻撃に対する国家安全保障を目的とするデータ保持および乗客名記録措置に関する欧州司法裁判所（CJEU）の判例により示唆された。CJEUによれば、これらの文脈においては、「国家安全保障に対する脅威は、現実かつ現在の、または少なくとも予見可能でなければならず、それは十分に具体的な状況が生じたことを前提とする」。[209]

防止

(348)　AI法第5条第1項(h)(i)および第5条第1項(h)(iii)と異なり、第5条第1項(h)(ii)にいうシナリオは、リアルタイムRBIの使用が具体的な人物の場所の特定または身元の特定のために許可されるとは明記していない。その目的は、特定の脅威の防止である。したがって、このシナリオでは、「移動中のテロリスト」、つまり、どこで行われるかは明確でなくても、人がテロリスト攻撃を計画しているとの具体的な兆候がある場合、同じ脅威に関連する複数の人物を探知しかつ追跡するためのリアルタイムRBIの使用も含まれ得る。

公園におけるテロリスト攻撃を防止するリアルタイムRBI 警察は、ある人物が、普段、テロリスト攻撃やテロリストグループに結び付く暴力的で過激なスローガンを叫び、公園周辺を走り回り、ナイフで攻撃する人々を探しているとの情報を得る。加盟国が、AI法第5条(1)(h)(ii)にいうシナリオにおいてリアルタイムRBIの使用を許可しているならば、法執行機関は、AI法第5条第2項ないし第7項の他の要件を満す場合、攻撃を防止するため公園内および公園周辺の人の身元を特定し、かつ場所を特定するためにリアルタイムRBIを使用することができる。

[205] https://cuta.belgium.be https://crisiscenter.be/en/risks-belgium/security-risks/terrorism-and-extremism
[206] https://www.government.nl/topics/counterterrorism-and-national-security/risk-of-an-attack-threat-level
[207] https://www.sgdsn.gouv.fr/vigipirate# https://www.sgdsn.gouv.fr/files/files/Vigipirate/20160130-np-sgdsn-pse-tackling-terrorism-together.pdf
[208] https://www.krisinformation.se/en/hazards-and-risks/terrorism
[209] 欧州司法裁判所2022年9月20日判決、SpaceNet、C-793/19（共同訴訟C-793/19、C-794/19）、ECLI:EU:C:2022:702、93項

9.3.4. 一定の犯罪の被疑者の所在の特定および身元の特定

(349) AI 法第 5 条(1)(h)(iii)は、「附属書Ⅱに定める犯罪であり、かつ、関係する加盟国において最長期間が少なくとも 4 年間である拘禁刑または留置命令によって処罰される犯罪に対する、刑事捜査、訴追または刑事罰の執行を目的とする、刑事犯罪を犯した被疑者の所在の特定または身元の特定」 のために公衆がアクセス可能な場所における RBI のリアルタイムの使用を認める。

AI 法附属書Ⅱは、上述の目的のためにリアルタイム RBI の使用が許可され得る重大犯罪を限定列挙する。これらの刑事犯罪は次のとおりである：

- テロリズム
- 人身売買
- 児童の性的搾取、児童ポルノ
- 麻薬または向精神薬の違法取引
- 武器、弾薬または爆発物の違法取引
- 殺人、重大な身体傷害
- 人間の臓器または組織の違法取引
- 核物質または放射性物質の違法取引
- 誘拐、違法監禁または人質にとること
- 国際刑事裁判所の管轄下にある犯罪
- 航空機または船舶のハイジャック
- 強制性交
- 環境犯罪
- 組織的なまたは武装した強盗
- 破壊工作
- 上記に列挙された犯罪の 1 つ以上に関与する犯罪組織への参加。

a) 所在の特定および身元の特定

(350) 加盟国は、犯罪捜査を行うため、犯罪を犯した者を訴追するため、または既存の判決を執行するため、刑事犯罪の被疑者の場所の特定および身元の特定のため、法の執行を目的として公衆がアクセス可能な場所内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) 被疑者および犯罪者

(351) AI 法第 5 条第 1 項(h)(iii)は、被疑者および犯罪者という 2 つのカテゴリーの個人を対象とする。被疑者とは、犯罪を犯したと信じるに足る充分な理由があり、かつ、その者が犯罪に関与していることを示す十分な証拠が既に収集されている者をいう。犯罪者とは、刑事犯罪を犯したとして訴追されまたは有罪判決を受けた者をいう。同じ条件（附属書Ⅱに列挙されている犯罪および最長期間が少なくとも 4 年の拘禁刑を科される犯罪）は、AI 法附属書Ⅱに列挙される犯罪の共犯者の場所を特定しまたは身元を特定するために適用される。

c) 重大犯罪のリスト

(352)　重大犯罪のみ、法の執行を目的とする公衆がアクセス可能な場所内においてリアルタイムRBIシステムを用することを正当化できる。

(353)　AI法附属書IIに列挙される最初の5つの犯罪は、TFEU第83条に列挙された「ユーロ犯罪」と同じであるが、他の犯罪は法執行協力のための優先事項となる。[210] それらのいくつか（たとえば、誘拐、核物質または放射性物質の違法取引）は、テロリズムに関連し得る。[211]

(354)　附属書IIに列挙されるすべての刑事犯罪は、被疑者または犯罪者に対する欧州逮捕状(EAW)の発付の要因となりうるが、これらの重大刑事犯罪の1つに被疑者の所在を特定しおよび身元を特定するためにリアルタイムRBIを使用することは、EAWの発付を要件としない。

(355)　さらに、この目的のためのリアルタイムRBIの使用は、それぞれの刑事犯罪が、最長期間を少なくとも4年とする拘禁刑の判決または留置命令により、関係加盟国において罰せられるものでなければならない。

都市における賑やかなフェスティバルの期間中、警察当局は、ライブ顔認識技術を導入し、フェスティバル周辺を監視し、違法な麻薬密売や性犯罪で未解決の逮捕状が発付された指名手配犯を特定する。フェスティバルのさまざまな入り口で、警察はモバイルカメラの前を通る人々のライブビデオ映像を使用し、指名手配犯の顔の監視リストとその顔を比較する。

まず、犯罪の種類については、RBIは、違法な麻薬取引の場合に使用することができる。しかし、性犯罪は、それが子どもの性的搾取、子どもの性的虐待素材、またはレイプに関連するものでない限り、犯罪のリストに含まれない。警察は、無差別で広範な方法で、すなわち、指名手配犯を見つけ街から排除することを期して、リアルタイム顔認識技術を導入することは許可されない。

警察が麻薬密売について欧州逮捕状の対象となる指名手配犯の身体的特徴を写真付きで受領し、指名手配犯がフェスティバルにいると警察が信じる理由がある場合には、状況は異なる。これらの状況において、リアルタイム顔認識技術を導入し対象の個人を識別することは、AI法第5条第1項(h)(iii)の対象となり得る。

クリスマスマーケットで12名が死亡する深刻なテロリスト攻撃の後、警察は、犯人を特定し、かつどこに犯人が逃げているのかを確認するため、リアルタイム顔認識技術を使用する。その状況においては、警察は、近くの鉄道駅および攻撃直後にそこから出発する列車の到着駅でリアルタイム顔認識技術を使用する。テロリスト攻撃の場合、そのような使用は、AI法第5条第1項(h)(iii)に基づき許可され得る。

[210] ユーロポールの優先事項。
[211] 前文33項参照、および指令2017/541第3条のテロリスト犯罪の定義参照。

(356) AI 法第 5 条第 1 項(h)(i)と第 5 条第 1 項(h)(iii)とは、AI 法第 5 条第 1 項(h)(i)に定めるシナリオの対象となる犯罪に関連づけることができる。リアルタイム RBI システムは、被害者または行方不明者を発見するために導入され得るが、人身売買、子どもに関する性的搾取（付属書 II に列挙されるとおり）、および誘拐（AI 法第 5 条 1 項(h)(i)にいう拉致が AI 法付属書 II に列挙される誘拐に該当する限り）の犯罪者または被疑者の所在を特定しかつ身元を特定するためにも使用され得る。また、AI 法第 5 条第 1 項(h)(ii)と(iii)とを関連づけることも可能である：リアルタイム RBI システムは、ある脅威が第 5 条第 1 項(h)(ii)の適用範囲となることを防止するために使用することができ、その脅威が現実となる場合、それらのシステムは「移動中」の犯罪者の身元特定/所在特定のために使用することができる。

10. 例外に関するセーフガードおよび要件（AI 法第 5 条第 2 項ないし第 7 項）

10.1. 対象者となる個人およびセーフガード（AI 法第 5 条第 2 項）

> AI 法第 5 条第 2 項は、次のように規定する。
>
> 第 1 項第 1 段落(h)に列挙する目的のいずれかを実現するための法の執行を目的として、公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用することは、具体的に対象となる人の身元の特定を確認するためにのみ当該(h)に定める目的で行われ、かつ、次の要素を考慮する：
>
> (a) 当該システム使用の可能性を生じさせる状況の性質。特に、当該システムが使用されなかったならば生じる害の重大性、蓋然性および規模；
>
> (b) 関係者すべての権利および自由に対する当該システムを使用することの結果。特に、当該結果の重大性、蓋然性および規模。
>
> 加えて、本条第 1 項第 1 段落(h)に列挙する目的のいずれかを実現するための法の執行を目的として、公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用することは、特に、期間的、地理的および人的制限に鑑み、この使用を認める国内法に従い、当該使用に関して、必要かつ相応なセーフガードおよび要件を遵守しなければならない。公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用することは、法執行機関が、第 27 条に従って基本的権利に対して与える影響の分析を完了し、かつ、第 49 条に定める EU データベースにこのシステムを登録した場合にのみ認められる。ただし、十分に正当化される緊急の場合において、この登録が不当に遅滞することなく実施されることを条件として、EU データベースへの登録なく、このシステムの使用を開始することができる。

(357) AI 法第 5 条第 1 項(h)(i)ないし(iii)に列挙された目的の 1 つに関するリアルタイム RBI システムの使用は、AI 法第 5 条第 2 項ないし第 5 条第 7 項に詳述される一定のセーフガードおよび要件に従う。

(358) まず、法の執行を目的として公衆がアクセス可能な場所内でリアルタイム RBI システムを使用することは、「具体的に対象となる人の身元の特定を確認する」ためだけに認められる。この最初の要件は、状況の重大性およびシステムを使用しないことによる損害と、技術が個人の権利およ

び自由に与える影響とのバランスをとることを目的とする。それは、リアルタイム RBI の導入のために、個人を対象とすることによる大規模な監視を回避することを目的とする。その結果として、法の執行を目的として公衆がアクセス可能な場所内でリアルタイム RBI システムを導入することは、対象となる個人に対してのみ許可されなければならない。

(359) 「身元の特定（confirming the identity）」という表現の使用は、「識別（identification）」というより、無差別な監視のリスクを制限する基本的権利に対する追加的なセーフガードとしての意味があり、前提として、AI 法第 5 条第 1 項(h)の意味における個人の識別が対象とされなければならない。この表現は、リアルタイム RBI の使用が、特定の個人が AI 法第 5 条第 1 項(h)(i)に列挙される犯罪の被害者であるか、または第 5 条第 1 項(h)(ii)もしくは第 5 条第 1 項(h)(iii)にいうシナリオのいずれかに関与していると法執行機関が信じる理由がある特定の個人もしくは情報が提供された特定の個人を捜索するためにのみ導入できることを意味するものとして理解されなければならない。これは、実際のところ、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との比較を意味する。AI 法第 5 条第 1 項(h)(ii)にいう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 条第 1 項(h)(iii)の意味における犯罪捜査の遂行については、法執行機関は、必ずしもシステムを使用する前に探している個人の身元を知る必要はない。法執行機関が特定の時間および場所において（誰が計画を実行するかを知ることなく）テロリストグループにより計画されたテロリスト攻撃に関する事実の兆候および情報を有する場合、法執行機関がテロリストグループの一部を構成する個人の生体データを含む参照データベースを構築しているのであれば、RBI システムは、テロリストグループから犯罪者を特定するために使用され得る。AI 法第 5 条第 1 項(h)(i)ないし(iii)にいう 3 つのシナリオすべてにおいて、「身元の特定」には、問題の人物の場所の特定も含まれ得る。

(360) 第2に、システムを使用する前に、システムの使用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であろう自然人、社会、法執行目的に対する害の重大性、蓋然性および規模が、関係者の権利および自由にシステムの使用が及ぼす影響、特に、それらの結果の重大性、蓋然性および規模に照らし、評価すべきである。これは、法執行機関またはその名で行動する主体が、より介入的でない代替の解決法が利用可能かどうかを評価することを含む。

たとえば、法執行機関は、一般的なセキュリティ、犯罪防止および過密状態の懸念に基づき、路上において、リアルタイム顔認識システムを使用することを禁止される。それはすべての人の絶え間ないモニタリングおよび監視を伴い、時間的に制限がなく、したがって、AI 法第 5 条第 1 項(h)に定める禁止の例外の基準に合致しないからである。

(361) 「**重大性**」の基準は、潜在的な害および結果に関連してここで適用されるが、危険にさらされている基本的権利への干渉の度合いにおける変化量を意味し、これは比例性の原則と関連する[212]。基本的権利への干渉について、ある干渉は他の干渉より深刻であると見なされる。

[212] 欧州司法裁判所 2018 年 10 月 2 日判決、Ministerio Fiscal、C-207/16、ECLI:EU:C:2018:788、55 項。ここで裁判所は「アクセスは、問題となる基本的権利への干渉の重大性に相応なものでなければならない」という。

(362) 「**規模**」の基準は、特に、干渉により影響を受ける人（子どもおよび脆弱な人または主流から外れた人を含む）の数とカテゴリーをいう。

(363) 最後に、「**蓋然性**」とは、ある事象が発生する可能性をいう。

(364) 害および結果の重大性、規模および蓋然性の評価は、すべて、法執行機関が履行する義務を負う基本的権利に対する影響評価の一部である（下記参照）。この評価は、ケースバイケースでの結論となる。

(365) 第3に、RBIのリアルタイムでの使用は、地理的範囲、期間、および対象者に関して明確に制限されなければならない。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ことを確保するためである。

(366) **地理的制限**に関しては、「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エリアを含み得る。生体識別の場合、これは、事象が発生する兆候があることにより、明確に線引きされた境界に対して地理的制限が適用されることを前提とする。このような線引きは、通常の状況下では、都市全体または国全体を含むべきでなく、より対象を絞らなければならない。

(367) 他のセーフガードは、措置の**人的範囲**に関するものである。すなわち**関係者のカテゴリー**を定義することである。これは、事件発生のさらなる兆候がない場合の、対象を絞らない無差別な人の識別を排除する。

(368) 最後に、**時間的期限**は、厳密に必要なものに限定された期間であるが、必要に応じ、適用されるルールに従って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、無期限または不確定な期間にすることはできない。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定める必要がある。

(369) 第4に、導入前に、リアルタイムRBIシステムを導入する法執行機関は、基本的権利に対する影響評価（FRIA）を実施し、当該システムをEUデータベースに登録しなければならない（十分に正当化される場合を除く）。

10.1.1. 基本的権利に対する影響評価

(370) AI法第5条第2項の適用により実施される基本的権利に対する影響評価（FRIA）は、AI法第27条に定める条件を遵守しなければならない。当該規定は、ハイリスクAIシステムに適用されるFRIAに関する要件を定める。

(371) AI法第5条の禁止事項が適用される（2025年2月2日以降）が、ハイリスクAIシステムに関する規定がまだ適用されない（2026年8月2日より前）間の期間においては、AI法第27条に定める基本的権利に対する影響評価（FRIA）の要件は、AI法第5条第1項(h)における1つ以上の例外から恩恵を受けるための条件を満たすリアルタイムRBIシステムの導入者により満たされ

なければならない。以下の暫定的なガイダンスは、ハイリスク AI システムの義務が適用される前の期間については、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI の使用のみに関連する。欧州委員会は FRIA のテンプレートを採択し、AI 法第 27 条に基づく義務に関するさらなるガイダンスを提供する。

(372) 基本的権利に対する影響評価（FRIA）は、RBI システムを含む、一定のハイリスク AI システムが基本的権利に与え得る影響を特定することを目的とする新型の影響評価である。FRIA は、説明責任のツールである。FRIA は、データ管理者（すなわち、個人データ処理責任者）が、LED 第 27 条、GDPR 第 35 条、または EUDPR 第 39 条に基づき実施しなければならない既存のデータ保護に対する影響評価（DPIA）に代わるものではない。

たとえば、DPIA は、生体データが、新たな技術を通じて処理され、公衆がアクセス可能な場所内において、自然人の権利および自由に対する高度なリスクをもたらす可能性がある場合（CCTV、AI 顔認識、身体装着型カメラなど)に、実施されなければならない。

(373) データ保護に対する影響評価（DPIA）は、個人データ処理が招来する個人の権利および自由に対するリスクに焦点を当てるが、基本的権利に対する影響評価（FRIA）は、AI システムが個人の基本的権利に与える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 の適用範囲は、対象となる活動および評価される基本的権利において、より広い。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、DPIA で既に取扱われている側面を対象とすることなく、かつ、重複を避け、データ管理者として、導入者が実施する DPIA を補完しなければならない[213]。これらのガイドラインにおける FRIA の分析は、リアルタイムでの RBI の許可された使用に限定され、AI オフィスがテンプレートを提供する前におけるこの暫定期間中に、導入者のための準備のガイダンスとなることを目的とする[214]。

(374) AI 法第 5 条第 2 項に基づく基本的権利に対する影響評価（FRIA）の実施義務は、RBI システムの導入者に課せられ、それらの代理として行動する主体もしくは組織またはいかなる者にも課せられない。他の行為者が導入者/法執行機関の代理として行動している場合、それらは、FRIA が適切に実行されることを確保するために、すべての関連する情報を用い FRIA の準備に寄与しなければならない。

(375) 基本的権利に対する影響評価（FRIA）は、許可されるリアルタイム RBI システムの導入前に、実施されなければならない。

(376) AI 法第 27 条により、基本的権利に対する影響評価（FRIA）は、以下の情報を含めなければならない。

・ 使用の意図目的を伴う、RBI 使用および導入者の使用プロセスの説明：

説明には、以下を含めなければならない。 - 導入者の名；

[213] AI 法第 27 条第 4 項
[214] したがって、当該分析は、一般的にハイリスク AI システムの場合を対象としない。

- リアルタイム RBI システムが使用されることになる法執行目的：
- それに対して生体識別が比較されることになる参照データベースの説明。これには、使用されることになる生体データの情報源（顔画像、音声サンプルなど）を含む；
- システムの機能を説明するためのシステムの基礎になる技術の説明（提供者が提供する利用可能な文書およびその名称を参照することによる）[215]；
- リアルタイム RBI が導入される法的根拠。

・ 使用期間および使用頻度

認められた例外の 1 つに対するリアルタイム RBI システムの個別の使用はそれぞれ、AI 法第 5 条第 3 項に基づく司法当局または他の独立の当局により、その導入前に許可されなければならない。これに対し、FRIA については、導入者は、意図された使用期間および予想される頻度の一般的な指標を提供しなければならない。

・ システムにより影響を受ける人およびグループのカテゴリー

AI 法第 5 条第 1 項(h)の例外について、FRIA は、以下を区別の目安としなければならない：
- 犯罪の被害者、犯罪者、または容疑者であり得る、対象となる個人、
- その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 システムが導入されることにな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。その生体データが比較のために使用される他の個人、通行人、および検索エリアに偶発的に存在した人々の権利もまた、影響を受ける。リアルタイム RBI システムが対象とする検索エリアの地理的範囲の説明は、システムが作用する人の数に影響する。

・ 影響を受けた人に対する害の具体的リスク:

(377) 法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI の使用により影響を受ける可能性のある基本的権利には、特に、次のものが含まれる。
- 私生活および家族生活に対する権利。公共の場における匿名性に対する人々の合理的な期待を含む；
- データ保護に対する権利。RBI システムは、特定の個人を識別するため、生体データおよびその他の個人データ（たとえば、名前、ID 番号、および民族などの機微データ）を処理することによる；
- 思想、良心および宗教の自由、表現の自由、ならびに探知される公共の場における集会および結社の自由。そこでは、個人が監視されていることを知っている場合、それらは行動を変えるか、一定の方法で行動することを自ら回避する可能性があるため、個人がその権利および自由を完全に行使することを避けることにより、RBI システムの使用が萎縮効果をもたらし得る；
- 効果的な救済および公正な裁判を受ける権利；

[215] ハイリスク AI システムに関するルールが効力を生じると、それは EU データベース内のシステムの登録番号およびそれ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実行し得る。

- システムが偏見（性別、民族、人種の偏見など）を埋め込み、かつ容疑者または犯罪者の誤認を導く場合の、差別を受けない権利；
- システムの対象にされていると感じることによる人間の尊厳に対する権利；
- 無罪の推定および弁護の権利；個人に悪影響を与えるいかなる決定も、リアルタイムRBIシステムの出力のみで行われ得ないことによる；
- 被害者、行方不明者、または容疑者が未成年者である場合の子どもの権利；
- 行方不明者の場合における高齢者の権利。

特定された影響を受ける人またはグループに影響を与える可能性のある害の特定のリスクを評価するため、基本的権利に対する影響評価（FRIA）は、影響を受ける可能性のある人を考慮に入れ、それらの人の基本的権利を特定し、影響の重大性およびその規模を含む、基本的権利に対する影響を評価しなければならない。

このFRIAは、また、その目的、およびより介入的でない代替手段の有無を含む、その使用が意図されている状況を考慮し、リアルタイム RBI システムの使用が必要かつ相応であるかどうかの評価を含むものでなければならない。FRIA は、技術文書に基づき、および利用できるのであれば、技術が偏見または差別を防止するためにテストされおよび開発されたトレーニングデータに基づき、システムの性能および精度の程度を説明しなければならない。

FRIA は、また、影響を受ける可能性のあるすべての個人の基本的権利、特に容疑者または犯罪者、捜索される被害者、および捜索の対象となる公衆がアクセス可能な場所に存在する他の個人の基本的権利に対し、リアルタイム RBI システムの使用が与える影響を特定しなければならない。システムがこれらの個人の生体データを処理する限り、その私生活および家族生活に対する権利およびデータ保護は影響を受けることになり、それはデータ処理活動に関する限り、データ保護に対する影響評価（DPIA）の一部として評価されることになる。リアルタイムRBIシステムの使用およびその他の基本的権利への影響に関連する他の活動について、FRIAは、DPIAを補完する。導入の状況により、これらの個人の他の基本的権利（たとえば、人間の尊厳に対する権利、思想、良心および宗教の自由、集会または表現の自由、効果的な救済および公正な裁判をうける権利、無罪の推定および弁護権、子どもの権利など）が影響を受ける可能性がある。

FRIAに基づく当該評価は、AIシステムの最初のサービス開始に先立ち、抽象的なレベルで実施されなければならない。リアルタイム RBI システムが使用される個別のケースそれぞれにおける使用の影響を決定する特定の状況次第となる考慮事項は、RBIシステムの各使用の司法当局または他の独立の行政当局による許可を求めるための個別の請求において、さらに詳述されなければならない（以下の10.23.8.3参照）。

・人間による管理措置

AI法第5条第3項によれば、リアルタイムRBIシステムの出力のみに基づき、個人に対して悪影響を与えるいかなる判断も下すことはできない。結果として、FRIA は、システムの動作中に追行することになる手続き、および意思決定プロセスの文脈において出力がどのように解

釈されるかを説明しなければならない。当該手続きは、RBI システムの導入に関する指示を提供し、出力の検証および解釈における人間のエージェントの役割を明確にし、かつそのシステムを操作するためのトレーニングを提供するものでなければならない。人間による管理を担当する者は、システムがどのように機能し、いつ性能が低下しまたは誤作動するかを理解するために、十分な「AI リテラシー、トレーニングおよび権限」[216]を有するものでなければならない。

AI 法第 14 条および第 26 条に基づく人間による管理およびモニタリングに関する他の考慮事項もまた関連性があり、かつ説明されなければならない。

・ リスク軽減措置

導入者は、人間による管理措置（差別的な措置を回避することを含む）を実施するだけでなく、ガバナンス手続きおよび不服メカニズム（たとえば誤認識の場合）を含む、リスクが具現化した場合の是正措置を説明しなければならない。

10.1.2. 許可された RBI システムの登録

(378) AI 法第 5 条第 2 項は、また、法の執行を目的とする公衆がアクセス可能な場所内で使用されるリアルタイム RBI システムの導入者に対し、AI 法第 49 条に定める EU データベースにシステムを登録することを義務づける。ただし、十分に正当化される緊急の場合（差し迫った脅威など）、法執行機関が不当に遅滞することなくシステムを登録することを条件として、登録前であっても着手できる。不当な遅滞とは、システムの使用前に、システムの登録を妨げた緊急の状況を考慮し、「できるだけ早く」という意味に理解されなければならない。登録がその基準を満たしているかどうかは、ケースバイケースでの評価が求められる。正確な時間的限界をあらかじめ定義することはできない。遅滞は、故意による行動によって引き起こされてはならない。AI 法第 49 条第 4 項によれば、法の執行を目的として使用される RBI システムは、制限された情報で、データベースのセキュアな非公開領域に登録され、かつその情報へのアクセスが制限される。

たとえば、法執行機関に対し、使用から 24 時間以内に RBI システムを登録することを要求することは、システムが生命に対する差し迫った脅威の状況において導入された場合、合理的な遅滞と見なされ得る。たとえば、銃乱射現場のシナリオにおける場合などである。

10.2. 事前の許可の必要性

(379) AI 法第 5 条第 3 項は、リアルタイム RBI システムの個別の使用それぞれについて、事前の許可を要求し、かつ**そのようなシステムの出力のみ**に基づく、不利な法的効果を生じさせる自動化された意思決定を禁止する。

AI 法第 5 条第 3 項は、次のように定める：

第 1 項第 1 段落(h)および第 2 項の目的による、公衆がアクセス可能な場所内における、法の執行を目的とするリアルタイム遠隔生体認識システムの各使用は、第 5 項に定める国内法の

[216] AI 法前文 91 項

詳細な規定に従って理由を示した請求に基づき発せられ、司法当局によってまたはこの使用が行われる加盟国をその決定が拘束する独立の行政当局によって付与される事前の許可に従う。ただし、十分に正当化される緊急の状況において、この許可が遅くとも 24 時間以内に不当に遅滞することなく請求されることを条件として、許可なく、当該システムの使用を開始することができる。当該許可が拒絶された場合、使用は即時に停止され、すべてのデータ、ならびに当該使用の結果および出力は、直ちに破棄されかつ削除されなければならない。

管轄を有する司法当局またはその決定に拘束力を有する独立の行政当局は、関係するリアルタイム遠隔生体認識システムの使用が、請求において示されたように、第 1 項第 1 段落(h)に列挙する目的のいずれかを実現するために、必要かつ相応であると当局が判断する場合にのみ、特に、当該使用が、時間的、ならびに地理的および人的範囲の観点から、必要最小限に限定される場合にのみ、客観的な証拠に基づいて、または当局に示された明確な指示に基づいて、許可を与える。当局が請求により決定する場合、この当局は、第 2 項に定める要素を考慮する。人に関して不利な法的効果を生じさせるいかなる判断も、リアルタイム遠隔生体認識システムの出力のみに基づき、下すことはできない。

10.2.1. 目的

(380) 法の執行を目的とする公衆がアクセス可能な場所内における「リアルタイム」RBI システムの使用につき、事前承認（「authorisation ex ante」）を要求する目的は、そのような目的によるそのようなシステムの想定される使用が、次のものかどうかに関し評価しおよび判断する必要による:
- 第 5 条第 1 項(h)(i)ないし(iii)に列挙する目的、すなわち、特定の被害者を対象とする捜索、特定の脅威の防止、または犯罪者の所在の特定または身元の特定の目的のいずれかを達成するために必要かつ相応であること；および
- 期間および地理的および人的範囲に関して厳密に必要なものに限定されること。

(381) これらの要件の結果として、法の執行を目的とする公衆がアクセス可能な場所内においてリアルタイム RBI システムを導入する前に、二重の必要性および相応性の評価が行われなければならない。第 1 に、基本的権利に対する影響評価（FRIA）を実施する場合、評価は、AI 法第 5 条第 2 項で求められるとおり、ユーザーにより行われなければならない。第 2 に、AI 法第 5 条第 3 項に従い、司法当局または独立の行政当局は、憲章および他の EU 法を考慮し、そのようなすべての使用の法的根拠を提供する国内法の範囲内において、そのようなシステムを使用する必要性および相応性も評価しなければならない。結果として、そのようなシステムはすべて、1)FRIA の後、および 2)管轄の国家当局がそのような使用を許可した場合のみ使用できる。

(382) AI 法第 5 条第 3 項は、AI 法第 5 条第 5 項と併せて読解され、かつ理解されなければならない：リアルタイム RBI システムの使用が許可されるためには、そのような使用を許可する当該加盟国において採択された国内法が存在しなければならない[217]。ある加盟国では、データ保護法のように、他の EU 法または国内法に基づき、生体システムの使用について、既に事前許可システムを導入している。

[217] AI 法第 5 条第 2 項も参照：「(...)この使用を認める国内法に従い(…)」.

10.2.2. 主要な原則: 司法当局または独立の行政当局による事前の許可

(383) AI 法第 5 条第 1 項(h)(i)ないし(iii)に列挙された目的のいずれかを遂行する関係する加盟国の国内法に定めるリアルタイム RBI システムの使用は、**その使用前に**司法当局または独立の行政当局によって許可されなければならない。これが主要な原則である。

(384) ただし、緊急の場合においては例外がある。これは十分に正当化されなければならない[218]。緊急性は、「問題となるシステムを使用する必要が、AI システムの使用を開始する前に許可を得ることを実質的かつ客観的に不可能とする状況」と説明される[219]。そのような緊急性のある場合、「AI システムの使用は、必要最小限に限定されなければならず、かつ、国内法において限定され、法執行機関自身により各緊急使用の文脈が特定されるような、適切なセーフガードおよび要件に従わなければならない」。

10.2.2.1. 国内の手続ルールに従う事前のかつ合理的な請求

a) 誰による請求か？

(385) 特定されていないが、請求は、通常、導入者、すなわち**所轄の（法執行）当局により**行われる想定され得る。AI 法第 3 条(45)b)に基づく法執行機関の定義によれば、「公共の安全に対する脅威からの保護および当該脅威の防止を含む、刑事犯罪の防止および探知、捜査または訴追または刑事罰の執行を目的として、加盟国の法が、公的機関の権限行使および公権力を委ねるその他のあらゆる組織または主体」が法執行機関とみなされ、事前の許可請求を提出するための「所轄当局」として責任者となる可能性もある。

(386) AI 法の適用範囲外となる活動のためのリアルタイム RBI システムの使用は、AI 法第 5 条第 3 項に基づき許可される必要はない。その後、そのようなシステムが法の執行を目的として使用されるならば、当該使用は、AI 法の適用範囲内となり、AI 法第 5 条第 1 項(h)の要件に合致する場合は許可が必要となる。

b) いかなる使用のための請求か？

(387) 法の執行を目的として公衆がアクセス可能な場所内で「リアルタイム」RBI システムを使用するには、たとえ当該システムが法執行機関の名で他の当事者、たとえばスポーツクラブやショッピングモールにより操作されている場合であっても、事前の許可が必要である。

たとえば： - 行方不明の子どもを捜索するための方策を委託された団体が、リアルタイム RBI システムの使用を決定する。それは、公的権限および公権力を行使し、犯罪を防止し、公共の安全に対す

[218]「法執行機関は、そのような状況においては、不当に遅滞することなく、遅くとも 24 時間以内に、より早い段階で許可を請求できなかった理由を示しつつ、その許可を請求しなければならない」ことを意味する。(AI 法前文 35 項)。
[219] AI 法前文 35 項。

る脅威を防止するための任務を遂行する権限がない。このような使用は、法の執行を目的とするものではないため、AI 法第 5 条第 1 項(h)に定める禁止事項に該当しない。しかし、このシステムは、「ハイリスク」（付属書 III の 1(a)）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の請求が必要となり得る。適用される国内法により、および GDPR 第 9 条第 1 項の例外のいずれかが適用されるかどうかにより、そのような処理に対する事前の承認も必要とされ得る。これに対し、同じ団体が法執行機関から、法執行の文脈において、管轄の法執行機関の監督および指示の下で、行方不明の子どもの捜索のために、その名で行動するように要請された場合、AI 法第 5 条第 3 項に従い、事前の承認が必要になる。

- 自然災害の犠牲者となるリスクのある人を支援するための手段の提供を委託された民間団体は[220]、その目的のためにリアルタイム RBI システムを使用することを決定する。このような使用は、法の執行を目的としたものではないため、AI 法第 5 条第 1 項(h)に定める禁止事項に該当しない。しかし、このシステムは「ハイリスク」（付属書 III の 1(a)）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の請求が必要となり得る。適用される国内法により、および GDPR 第 9 条第 1 項の例外のいずれかが適用されるかどうかにより、そのような処理に対して事前の承認も必要とされ得る。

c）「各使用」はいつか？

(388) AI 法第 5 条第 3 項に従い、「各使用」に事前の許可が必要である。これは、そのような許可を取得する確定的な時が、リアルタイム RBI システムをインストールする前の時ではなく、その各具体的な使用の時であることを前提とする。

たとえば：
- 警察は、都市の主要鉄道駅に生体対応の CCTV カメラを設置する（AI 法に基づく許可は必要でないが、生体システムはハイリスクシステムの要件に従わなければならない。基本的権利に対する影響評価（FRIA）は最初の使用前に準備されなければならず、システムの各個別使用前に司法当局または独立の行政当局による個別の許可が必要である）。

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有する（リアルタイムの識別には事前の許可が必要である）。

d）理由を示した請求

(389) AI 法第 5 条第 3 項は、リアルタイム RBI の使用に対する各個別請求が「根拠がある」ものであり、したがって、実証されかつ理由が示されることを求める。

(390) 一部の加盟国は、そのような請求をオンラインで提出することを認める[221]。AI 法第 5 条第 5 項に従い、国内法は、厳密な請求内容に関する要件を定めなければならない。これには、上記の概

[220] 自然災害には、川の氾濫や自然発火などを含む。
[221] たとえば、フランスのデータ保護機関である CNIL への許可請求を参照。

略的要件を十分に考慮し、リアルタイム RBI の使用のために厳密な必要性と相応性を決定する十分な証拠、およびそのような使用を許可する例外的な性質を反映する他の関連する側面を含む。

10.2.2.2. 司法当局または独立の行政当局による許可

(391) 当該許可は、その決定が拘束力を有する司法当局または独立の行政当局によってのみ与えられる。

a) 独立の権限

(392) 欧州司法裁判所（CJEU）は、さまざまな文脈において、「独立」の概念を解釈してきた。たとえば、HK v Prokuratuur 事件における CJEU の説明によれば、独立とは、当局が「中立の立場」を維持することを意味する[222]。CJEU が明示するところでは、以前の調査に関与した当局（この場合は検察官である）に、そのような独立性はない。同様の考察は、AI 法第 5 条第 3 項により求められる独立性に関しても適用される可能性があり、許可を与える当局は、前提として、RBI システムを使用する権限から独立でなければならない。これは、警察だけでなく、警察の職務および許可が求められている RBI の使用を監督する予審判事または検察官の場合にも適用される。

(393) Commission v Poland 事件は、鉄道の安全性の文脈において、組織が独立しているとみなされる場合を問題として扱った事件である。欧州司法裁判所は、「公的組織に関し、独立性は、通常、問題の組織が指示や圧力から保護され、その独立性が確保されるべき組織との関係において、完全に自由に行動できることを確保する地位をいう」と判断した[223]。同様の指摘は、AI 法第 5 条第 3 項の文脈で適用され得る。

(394) 民主主義社会における司法当局は、一般に独立の当局でもある。司法は、行政府や立法者から独立している場合に重要な役割を果たし、法の適用ならびに基本的権利および自由を自律的かつ独立の方法で扱い、見直す。司法の独立は、法の支配の重要な側面の一つであり、（憲章）第 47 条および ECHR 第 6 条 1 項によって保証される[224]。

b) 利用が行われる場所の権限

(395) 許可は、国内法に従って、管轄を有する当局宛としなければならない[225]。

c) 許可は、例外に定める目的のいずれかを達成するため「必要かつ相応な」場合に限る

[222] 欧州司法裁判所 2021 年 3 月 2 日判決、Prokuratuur、C-746/18、ECLI:EU:C:2021:152、54 項。
[223] 欧州司法裁判所 2018 年 6 月 13 日判決、Commission v Poland、C-530/16、ECLI:EU:C:2018:430、67 項。
[224] R. Manko, Judicial Independence in the case law of the European Court of Human Rights, 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, 12 p.参照
[225] 欧州司法裁判所 2015 年 10 月 6 日判決、Schrems, C-362/14, https://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:62014CJ0362 , ECLI:EU:C:2015:650, 44 項参照。

(396) 法の執行を目的として公衆がアクセス可能な場所内においてリアルタイム RBI を使用するいかなる許可も、AI 法第 5 条第 3 項の要件に合致するかどうか評価しなければならない。

高度に介入的

(397) データ保護の文脈において、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)によりそのガイドライン 5/2022 において、および欧州データ保護監督機関(EDPS)により、いくつかの基本的権利および自由に影響を与えるとみなされる。当該見解は、EU 基本権機関および欧州評議会によって共有される[226]。欧州司法裁判所（CJEU）[227]および欧州人権裁判所（ECtHR）[228]の双方が、生体データ処理の機微性を認める。

(398) **基本的権利および自由における**いかなる干渉も、**常に権利および自由の本質を尊重しなければならない**。これは、憲章第 52 条第 1 項に従うものである。

(399) 基本的権利および自由の「本質」の概念は、欧州司法裁判所（CJEU）の判例法において発展してきた、EU の法秩序における独立の価値である。基本的権利または自由の本質が尊重されない場合、それは、ある措置により権利または自由が不当に侵害され、事前にいかなる干渉も許されないことを意味する。

「必要かつ相応な場合」のみ

(400) 基本的権利および自由に対するいかなる干渉も、原則として、憲章第 52 条に従い必要性および相応性を尊重する「法」を必要とする（以下の AI 法第 5 条第 5 項参照)。AI 法第 5 条第 3 項は、次のことを求める。すなわち、法の執行を目的とする公衆がアクセス可能な場所内におけるリアルタイム RBI の使用を許可する国内法は、AI 法第 5 条第 1 項 (h)に「列挙する目的のいずれかを実現するために、必要かつ相応であると当局が判断する場合にのみ」そのような使用の許可が認められることを規定しなければならない。国内当局は、生体識別が厳密に必要かどうか検証しなければならない[229]。この評価は、特定の状況における各使用の許可を請求する前に、一般的な方法により必要性および相応性の評価を既に含むべき FRIA（基本的権利に対する影響評価）に基づくものでなければならない。

10.2.2.3. 事前許可の要件の例外：24 時間以内の請求および拒否された場合の結果

(401) 緊急の場合、ユーザーは、リアルタイム RBI システムが使用された時から 24 時間以内に許可の請求を提出できる。実際には、一般に、生体対応カメラまたは生体適応カメラが「オン」のスイッチが入りおよび導入された時で、かつ最初の生体比較がシステムにより行われた時である。

---

[226] 個人データの自動処理に関する個人の保護のための協定の諮問委員会(ETS108)、顔認識ガイドライン、2021 年。
[227] 欧州司法裁判所 2023 年 1 月 26 日判決、Ministerstvo na vatreshnite raboti、C-205/21、ECLI:EU:C:2023:49、60 ないし 76 項、および 116 ないし 134 項。
[228] 欧州人権裁判所 2023 年 7 月 4 日判決、Glukhin v Russia、申立番号 11519/20、ECLI:CE:ECHR:2023:0704JUD001151920、88 項および 90 項（以下、「グルヒン対ロシア判決」という）。
[229] データ収集については、司法裁判所 2024 年 11 月 28 日判決、Ministerstvo na vatreshnite raboti、C-80/23、ECLI:EU:C:2024:991 も参照。

処理活動のログは、当局が、請求のタイムラインを実証するために利用できるようにしなければならない[230]。

(402) このような場合、請求には、システムの使用開始前に、事前の請求が提出されなかった理由を示さなければならない。

10.2.2.4. 許可請求が拒否された場合の即時停止およびデータ削除

(403) AI 法第 5 条第 3 項は、さらに、緊急の場合に許可請求が拒否された場合、リアルタイム RBI システムの使用を直ちに中止しなければならないと規定する。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、直ちに破棄されかつ削除されなければならない。[231] AI 法第 5 条第 3 項は、例外なく、この点に関して明確である。導入者は、以下を有する：

a) 生体情報（たとえば、顔画像、音声スニペット）および適宜、関連する識別情報を含む参照データベース。これは、
b) その個人を特定し選び出すために、公衆がアクセス可能な場所内に存在する個人から取得された生体情報と比較され、
c) この照合により比較結果を得る。

(404) 収集および処理されたデータを破棄しおよび削除するとの要件は、許可されない生体識別のために使用された参照データベースが問題の捜査のために特別に構築された場合、それが除去されおよび削除されなければならないことも意味する。法執行機関が構築し、かつリアルタイム RBI の許可されない使用*以外*の正当な目的のために、*合法的な方法*で、識別に使用されるデータベースを維持することを意図していた場合*のみ*、データベースを維持することができる。

(405) 生体情報を含む、(違法な) データベースの削除に加え、収集されたすべての画像およびその他の個人データ（メタデータを含む)、技術処理データ（テンプレートおよびその他の個人データを含む)、およびリアルタイム RBI システムの違法な使用中に取得されたその他の比較および出力データも、削除されなければならない。

(406) 法執行機関が拒否に異議を申立てる場合、その申立てについて終局決定が下されるまで、受託者はデータを保持することができる。その期間中、これらのデータは、一般に、法執行機関の処分に供されてはならない。[310] (77)

10.2.2.5. リアルタイム RBI システムの出力のみにより判断しないこと

[230] 自動生成されたデータログは、ハイリスク 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ない。さらに、附属書 III の 1(a)にいうハイリスクシステムについては、各使用につき開始データと終了データ、および機関を含めなければならない。AI 法第 12 条第 3 項(a)および第 19 条参照。
[231] 監督当局は、この事後的ファクトチェックおよび管理を行う権限をも有するものでなければならない。AI 法第 5 条第 5 項参照。

(407)　AI 法第 5 条第 3 項に従い、リアルタイム RBI システムの導入者が許可を得たとしても、「リアルタイム」RBI システムの出力のみに基づき、人に不利な法的効果をもたらす判断を下すことはできない。

> たとえば：
> - 人が、顔認識システムによる識別のみに基づき、それ以上の検証がないまま、重大な犯罪により逮捕され、そして投獄される。これは、AI 法第 14 条に基づく人間による管理に関する要件に加わるものである。検証は、たとえば、特定の人が別の場所にいたかどうか、またはその人が捜査対象者となり得ない他の理由があるかどうかとの問題に関係する。

人間による管理に関する AI 法第 14 条の要件

(408)　AI 法第 5 条第 1 項(h)に列挙される目的の 1 つを追求し、かつ第 5 条第 2 項ないし第 6 項を遵守することにより許可されたリアルタイム RBI の使用は、依然としてハイリスクシステムのルールの適用を受ける。AI 法第 14 条に従い、ハイリスク AI システムは、その「設計および開発が、その使用期間中、特に、適切なヒューマン・マシン・インターフェース・ツールにより、自然人による効果的な管理を可能とするものでなければならない」。AI 法第 14 条第 5 項に従い、導入者は、「必要な能力を有し、訓練を受け、かつ権限を有する少なくとも 2 名の自然人によって、その特定事項がそれぞれに検証されかつ確認されない限り」、または 「EU 法または国内法が、当該要件の適用が不相応であると判断する」場合を除き、システムからもたらされる特定事項に基づき、いかなる措置も判断も行うことはできない。AI 法第 4 条は、AI システムが使用されることになる人を考慮し、「AI システムの操作および使用に従事するそれらの被用者およびその他の者について、十分なレベルの AI リテラシー」を確保するため、AI システムの提供者およびユーザーの AI リテラシー対策を規定する。

(409)　データ保護の文脈において欧州データ保護委員会が述べるように、人間による管理が効果的であるために重要となるのは、「(この場合は顔認識) システムおよびその限界を理解し、かつ、その結果を適切に解釈できるようにすることである。また、自動化バイアスの影響を打ち消し、たとえば、時間的制約、煩雑な手続き、キャリアへの潜在的な不利益によって、結果を無批判に受容することを助長しないようにする職場や組織を確立することが必要である」[232]。 AI 法の文脈においても同様の考察が適用され得る。

10.3. 法の執行を目的とする公衆がアクセス可能な場所内における「リアルタイム」遠隔生体認識システムの各使用に関する当局への通知

> AI 法第 5 条第 4 項は、次のように規定する。
> 第 3 項を害することなく、法の執行を目的とする、公衆がアクセス可能な場所内での、リアルタイム遠隔生体識別システムの各使用は、第 5 項に定める国内規定に従って、関係する市

[232] 欧州データ保護委員会 (EDPB)、法執行分野における顔認識技術の使用に関する 2022 年 5 月のガイドライン 2.0 版、2023 年 4 月 26 日、22 頁。

場監視当局および国内のデータ保護当局に通知される。この通知は、少なくとも、第6項に定める情報を含むが、機微取扱データを含まない。

(410) AI法第5条第1項(h)(i)ないし(iii)に列挙される目的の1つを追求するRBIシステムの各使用は、関連する市場監視当局および国内データ保護当局に通知されなければならない。通知は、許可件数およびその結果について報告できるようにするため、各使用の後に行わなければならない。通知には、機微取扱データを含める必要はない。AI法第3条(38)によれば、「機微取扱いデータ」とは、法執行活動（犯罪の防止、探知、捜査、または訴追）に関連する取扱データであって、その暴露が刑事手続の完全性を損ね得るものをいう。

(411) 報告に関する要件の詳細については、以下の10.6を参照。

10.4. AI法の例外の範囲内における国内法の必要性

10.4.1. 原則：全部または一部の例外の許可に関し法的根拠を提供するために必要な国内法

(412) 法の執行を目的として、公衆がアクセス可能な場所内で、「リアルタイム」RBIシステムの使用の運用を開始するには、国内法が必要である。同時に、AI法第5条第5項は、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と規定する。AI法は、リアルタイムRBIの使用を認める国内法が採択された場合、AI法に定める要件を遵守するために国内法に定めなければならない実質的要素を明記する。

AI法第5条第5項

加盟国は、第1項第1段落(h)および第2項・第3項に列挙する範囲内および条件下において、法の執行を目的として、公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用することについて、全面的または部分的に許可する可能性を規定することを決定することができる。関係する加盟国は、第3項に定める許可の請求、許可の付与および実施に必要となる、ならびにこれに関する管理および報告に必要となる詳細な規定を、その国内法において定める。また、これらの規定は、管轄当局が、第1項第1段落(h)に列挙するどの目的について、および、特に、(h)の(iii)に定めるどの刑事犯罪について、法の執行を目的として、当該システムを使用する許可を与え得るかを定める。加盟国は、その採択後、遅くとも30日以内に、欧州委員会に対し、これらの規定を通知する。加盟国は、EU法に従って、遠隔生体識別システムの使用に関し、より制限的な法を採択することができる。

10.4.2. 国内法は、AI法第5条第1項(h)の制限および要件を遵守しなければならない。

(413) 法の執行を目的とする公衆がアクセス可能な場所内での「リアルタイム」RBIシステムの使用は基本的権利への干渉とみなされるため、AI法第5条第5項は、そのような使用が加盟国の国内法によって規定されなければならないことを定める。これらの国内法は、そのようなシステムの使用に関する法的根拠を定める。

(414) 国内法は、AI 法第 5 条第 1 項(h)に定める制限を超えてはならないし、AI 法に定めるその他のすべての関連条件を尊重しなければならない。これは、加盟国が、AI 法第 5 条第 1 項(h)(i)ないし(iii)に列挙されるものを超えて、法の執行を目的とする公衆がアクセス可能な場所内でリアルタイム RBI を使用できる目的を拡張できないことを意味する[233]。

(415) 加盟国は、その国内法の採択後、遅くとも 30 日以内に、欧州委員会に対し、その国内法を通知しなければならない。当該通知は、加盟国の法が AI 法に準拠していることを推定するものではない。AI オフィスは、通知受領後、受領確認を送付する。また、加盟国は、その採択に先立ち、提案された国内（または地域の）法の暫定版を、AI オフィスに送付することも奨励される。いずれにせよ、第 5 条第 5 項に定める採択後 30 日の法定期限内に AI オフィスに通知しない場合、他の文脈[234]においても判断されているように、国内法は法的手続き上執行不能となる可能性がある。欧州委員会は、加盟国の法を公開のウェブサイト上で公表する。

(416) 加盟国は、EU 法に従い、より制限的な法律、すなわち、AI 法第 5 条第 1 項(h)および第 2 項ないし第 7 項に定める要件よりも厳しい法を導入することができる。

10.4.3. 許可の請求、付与および実施に関する詳細な国内法

(417) 許可の請求、付与および実施に適用される詳細なルールについては、国内法により定められることになる。問題となるシステムの使用を許可しようとする各加盟国は、その国内法において、当該ルールを明記しなければならない。それらのルールは、そのような使用の厳密な必要性および相応性について判断できるよう、リアルタイム RBI システムの使用に関する関連性のある完全な情報を、許可を与える当局に対し提供することを目的とする。

リアルタイム RBI システムの使用を許可する国内法は、たとえば、次のように規制し得る。

- AI 法第 5 条第 1 項(h)の対象となる管轄当局、および許可を付与（または拒絶）する権限を有する加盟国の独立当局となる者；

- 公衆がアクセス可能な場所内におけるリアルタイム RBI が法の執行を目的として使用される可能性のある目的の詳細な範囲（第 5 条第 1 項(h)(i)ないし(iii)に列挙される目的を超えることなく、それらをさらに限定することは可能である）；

- 請求は書面で行い、その使用を正当化する具体的な犯罪／状況について具体的な使用および使用の意図目的の詳細な説明を要求することを規定すること；

[233] 欧州司法裁判所 2022 年 4 月 5 日判決、An Garda Síochána コミッショナー、C-140/20、ECLI:EU:C:2022:258、54 項参照：「相応性の要件を満たすため、国内法は、問題の措置の範囲および適用を定め、ならびに最小限の保護措置を課す、明確かつ詳細なルールを定めなければならない。」

[234] 類推による。欧州司法裁判所 2019 年 12 月 19 日判決、Airbnb アイルランド、C-90/18、EU:C:2019:1112、96 から 97 項参照。

-特に場所、期間、および人的範囲に関し、AI 法第 5 条第 1 項(h)(i)ないし(iii)に列挙された目的を追求するシステム使用を正当化するため、ならびにシステム使用の関連性、十分性と有効性、およびより介入的でない手段がないことを含む、厳密な、必要性と相応性を正当化するための理由を要件とし、および裏付けとなる証拠を提出すること（および適宜、翻訳を必要とすること）；

- 使用される技術およびデータ収集の地点の説明；

- 使用されるシステムの最低限の信頼性、使用される閾値、および精度の率；

- 許可を与える当局が、技術的な詳細および精度の基準を含む提出された情報を、事前および事後にいつでも監査する可能性；

- 使用される参照データベースの仕様；

- 取得されたデータおよび使用された他のあらゆる関連の個人データの保存期間；

- データへの不正アクセス防止を含む、セキュリティ対策；

- その他の保護措置（必要に応じて）；

- 他国を含む、民間または公的機関との協力ならびにデータ移転および交換の説明；

- プロセスのトレーサビリティ；

- 導入者の責任者の名；

その他の形式的要素に関する発出について

- 聴聞により補完される書面による手続きの可能性；

- 拒絶の理由;；

- 捜査される者の権利、データが取得される者の権利、および第三者の潜在的権利[235]；

- 当局がその決定を下すべき期間；

- 許可の付与／拒絶の際の正式な通知の必要性；

[235] たとえば、EDPB、法執行分野における顔認識技術の使用に関する 2022 年 5 月のガイドライン 2.0 版、2023 年 4 月 26 日、24 頁以下参照。

- (形式的および実質的) 要件の不遵守に対する制裁；

- 許可が拒絶された場合の不服申立ての権利；

実施について

- 実質的要素の概要とともに、中央レジスタにリアルタイム RBI システムの使用を登録すること；

- 追加的報告義務の可能性；

- 許可を延長または変更する手続。

10.4.4. 許可に対する監督および報告に関する詳細な国内法

(418) AI 法第 70 条は、加盟国に「少なくとも一の認定当局および少なくとも一の市場監視当局を設置する」ことを義務づける。AI 法第 74 条第 8 項は、「加盟国は、規則(EU) 2016/679 に基づくデータ保護の監督に関する管轄当局か、指令(EU)2016/680 第 41 条ないし第 44 条に定める同じ条件に従って指定された他のあらゆる当局のいずれかを、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。」と規定する。

(419) これは、許可を与える当局の指定に加えて行われるものであり、加盟国は、AI 法第 5 条第 1 項 1(h)(i)ないし(iii)に列挙される目的のいずれかについてリアルタイム RBI システムの使用を許可する前に設置しなければならない。

10.5. 加盟国の国内の市場監視当局および国内のデータ保護当局による年次報告書

AI 法第 5 条第 6 項は、次のように規定する。
第 4 項に従って、法の執行を目的として、公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用することを通知された加盟国の国内の市場監視当局および国内のデータ保護当局は、この使用に関する年次報告書を欧州委員会に提出する。この目的のため、欧州委員会は、加盟国ならびに国内の市場監視当局および国内のデータ保護当局に対し、テンプレートを提供しなければならず、これには、管轄を有する司法当局によって下された、またはその決定が第 3 項に基づく許可の請求に拘束力を有する独立の行政当局によって下された決定の数に関する情報、およびその結果に関する情報を含む。

(420) 法の執行を目的とする公衆がアクセス可能な場所内でのリアルタイム RBI システムの使用について、導入者から通知された加盟国の国内の市場監視当局および国内のデータ保護当局(第 5 条

第4項参照)は、欧州委員会に対し、当該使用に関する年次報告書を提出しなければならない。これらの報告書は、欧州委員会が提供するテンプレートに基づき作成される。このテンプレートは、しかるべき時期に策定される。

(421) 導入者がEUの機関、組織、または事務所である場合、欧州データ保護監督官が、法の執行を目的として公衆がアクセス可能な場所内で使用されるリアルタイム RBI システムについて、毎年、欧州委員会に通知する義務を負う。

(422) AI法は、加盟国に対し、2025年2月2日より前に国内の市場監視当局を任命することを要求していないため、2025年2月2日から2025年8月2日までの期間を対象とするのは、国内のデータ保護当局の報告書のみである。

(423) 国内の市場監視当局および国内のデータ保護当局は、個別の報告書を提出するか、加盟国ごとに共同の報告書を提出するかを自由に決定できる。

10.6. 欧州委員会による年次報告書

> AI法第5条第7項は、次のように規定する。
> 欧州委員会は、第6項に定める年次報告書に基づき、加盟国において収集されるデータに立脚して、法の執行を目的として、公衆がアクセス可能な場所内で、リアルタイム遠隔生体識別システムを使用することに関する年次報告書を公表する。これらの年次報告書は、関連する法執行行為における機微取扱データを含まない。

(424) AI法は、欧州委員会に対し、加盟国およびEUの機関、事務所、および部署による法の執行を目的とする公衆によりアクセス可能な場所におけるリアルタイム RBI システムの使用に関する年次報告書を、集計データに基づき公表することを求める。これらの報告書は、AI法第5条第6項に従い、各国の当局から通知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は、機微取扱データを含めないこととする。機微取扱データとは、「刑事犯罪の防止および探知、捜査、または訴追の各活動に関連する取扱データであって、その暴露が刑事手続きの完全性を損ね得るもの」をいう[236]。これが意味し得るところは、たとえば場所、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は公表されないことである。

10.7. 適用範囲外となるもの

(426) AI法第5条第1項(h)による禁止事項の対象とならない RBI システムのその他すべての使用は、AI法の範囲内にある場合に限り、AI法第6条で定められ、かつ付属書III、1(a)に列挙されるとおり、ハイリスクAIシステムの範疇に入る。

[236] AI法第3条(38)。

(427) AI 法第 5 条第 1 項(h)の禁止事項の範囲外にある RBI システムは、生体確認/認証システムを含み、および法の執行を目的とする公衆がアクセス可能な場所における（事後的）RBI システムの遡及的使用を含む。たとえば、警察当局は、犯罪容疑者の画像を犯罪データベースに記録された顔画像と比較するため、遡及的な顔認識を実施することを、国内法によって許可される場合がある[237]。禁止事項の範囲外となる他の使用は、私的空間（誰かの家など）またはオンライン空間（子どもの性的虐待素材を拡散した被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使用など）のいずれかにおいて、法の執行を目的としてリアルタイム RBI システムを使用することである。最後に、民間の行為者による RBI システムの使用は、リアルタイムおよび遡及のいずれも、禁止の範囲外である（スーパーマーケットが既知の万引き犯を特定するためライブ顔認識技術を使用する、スポーツアリーナがアリーナへの入場を禁止された個人を特定するためライブ顔認識技術を使用する、または学校においてセキュリティ目的や出席確認のためライブ顔認識技術を使用するなど)。

(428) ハイリスク AI システムに一般的に適用されるルールに加え、法の執行を目的とする RBI システムの遡及的使用は、AI 法第 26 条第 10 項により追加的要件および保護措置の対象となる (2026 年 8 月 2 日から適用)。[238]

(429) **法の執行以外の目的**での使用は、いかなる場合にも、**データ保護ルール**を遵守しなければならない。以下のケースは、このような使用の場合における GDPR 第 9 条第 2 項の解釈、および生体データの処理の例外を示すものである。

たとえば

- フランスの行政裁判所は、アクセス管理およびセキュリティの目的で、2 つの公立学校でライブ顔認識技術を試験的に実施したことについて、（データ保護ルールの下において）必要性もなく相応でもないと判断した。たとえば、バッジの使用など、学生にとってより介入的でない代替手段が利用できた。加えて、明示的な同意の要件を満たしていなかった。したがって、同意は、高校において、顔認識技術を試験的に実施するための有効な法的根拠として用いることはできなかった。 [239]

- オランダでは、スーパーマーケットが、万引き防止のためにライブ顔認識技術を使用することが認められなかった。顧客からの明示の同意、または具体的な公共の利益のため（セキュリティ目的など）の処理を許可する法的根拠がなければ、スーパーマーケットは、生体データを処理し、したがって顔認識技術を導入することができなかった。[240]

[237] たとえば、フランスのデータベース Traitement des Antécédents Judiciaires。Décret no. 2012-652 du 4 mai 2012 relatif au Traitement des Antécédents Judiciaires (Decree 2012-652) により創設。
[238] AI 法第 26 条第 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-recognition-technology

- フランスでは、サポーターを識別するために、サッカークラブの入り口でライブ顔認識技術を使用することが禁止され[241]、スペインでは、観客の安全を確保するための使用が禁止された。[242]

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入り口付近の警察車両に、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、区域の安全を確保し、指名手配犯の特別監視リストのデータベースにその顔が記録されている個人を特定する。この監視リストには、犯罪（重大犯罪から詐欺や強盗まで）を犯したことが疑われる者、諜報目的により潜在的関心の対象となる者、精神的に問題がある脆弱な者が含まれる。警察によるライブ顔認識技術の使用は、イベントに特定の人物が存在していることに関する情報とは関連しない。リアルタイム RBI の使用が認められ得る者を捜査するための監視リストに、その者が存在する可能性はあるとしても、このリストはあまりにも具体的でなく、サッカーの試合というイベントとは関連しない。したがって、このような使用は禁止される。

生体識別システム（遠隔でない）は、人が、原子力発電所にアクセスできるかどうかを確認する。人がカメラ（明白であるもの）の前に現れ、アクセスがシステムにより拒絶された場合、当該システムは、その後、その者がテロリストの監視リストに存在するかどうかを識別しようとする。当該システムは遠隔ではない。人は、発電所への立入り許可を得るための認証作業に積極的に参加していた。このユースケースは、AI 法第 5 条の禁止事項に該当しない。

大都市の警察当局は、ライブ顔認識技術を実施できる AI 搭載の CCTV カメラを導入する。おそらく、顔認識に加え、物体検知や群衆の動きなど、さまざまな機能が追加される可能性がある。

これらのカメラは、礼拝所、LGBT+コミュニティがよく集まる多くの場所、診療所、薬局、さまざまなレストランやバーなど、複数の場所に設置される。

生体認識対応カメラの設置自体は、AI 法で禁止されない。

ただし、不特定かつ無差別な自然人の認識を含む、一定の使用は禁止される。

夏季休暇中、住宅街において、数件の空き巣被害が発生した。

警察は、空き巣被害の前に、近隣で異なる時期に、被疑者を目撃した者から被疑者に関する証言を入手した。被疑者を特定し逮捕するために、警察は、週末に、近隣の各地でライブ顔認識技術を使用する。目撃者の証言に基づき、警察は、容疑者の合成写真を作成し、管理データベースから合成写真に似た個人の写真をいくつか抽出した。

たとえ警察が対象である被疑者に対しライブ顔認識技術を使用し、使用範囲と時間を定めていたとしても、AI 法付属書 II に列挙されていない犯罪の場合、当該使用を導入することは認められない。

[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

警察は、生体認識システムを使用し、サッカースタジアムにおいて、ファンの感情をスクリーニングする。このシステムは、潜在的な攻撃を発見し、スタジアムの当該部分に、リアルタイム RBI を即時に導入し、過去に暴力的であったフーリガンを特定する。

スタジアムにおける感情のスクリーニングは、AI 法のもとで禁止されない（依然として AI 法のハイリスクの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、法の執行を目的として人を特定する必要性について決定するのが生体認識システムである場合、AI 法に基づき禁止される。

警察は、市街地および地下鉄に設置された CCTV ネットワークを利用し、街頭で集団抗議活動を組織した政治に対する抗議者を特定する。関係加盟国において、道路などの公道および公共エリアにおいて行われる集団抗議活動の主催者は、公共の混乱や暴力を防ぐため、計画された抗議行動の 3 日前までに地方自治体当局に通知しなければならない。通知がない場合、6 ヶ月以下の拘禁刑および 8,000 ユーロ以下の罰金が科され得る刑事犯罪となる。抗議者を特定するため、警察は街頭に設置された CCTV カメラから映像を抽出し、抽出した画像とソーシャルメディアに投稿された写真とを照合し、遡及的な顔認識を実施する。

顔認識技術の遡及的使用は、AI 法により禁止されない。その使用は、ハイリスクとみなされ、そのようなシステムに関する AI 法の要件を遵守しなければならない[243]。

その他の禁止されない行為の例：

- ホテルがリアルタイム RBI を使用し VIP の来客を認識する場合。これは法の執行ではない。
- ショッピングモールが万引き犯を見つけるためにリアルタイム RBI を使用する場合。これは法の執行ではない。

禁止される場合：

警察から委託され、ショッピングモールが、リアルタイム RBI を使用して万引き犯を発見する。このシステムは、法の執行を目的として公衆がアクセス可能な場所に導入される。万引き犯の捜査は AI 法第 5 条第 1 項(h)の例外のいずれにも該当しないため、この利用は禁止される。

11. 適用開始

[243] 法の執行を目的とする生体データの処理は、引き続き LED 第 10 条の適用対象であり、それは国レベルで実施される必要がある。FRT（顔認識技術）の事後的使用を実行するためのその処理は、それが厳密に必要である場合にのみ認められ、かつ適切な保護措置の対象とされなければならない。デモ参加者を特定するために、FRT の事後的使用が、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 Glukhin 対ロシア判決において、欧州人権裁判所は、犯罪の探知は正当な目的となり得るとしても、事後的およびライブ双方での FRT の使用は、公共の秩序または輸送の安全性に対するリスクがないため、不相応であると判断した。裁判所は、FRT の「非常に介入的な」性質を強調した。本件において、裁判所は、FRT の使用は差し迫った社会的必要に応えるものではなく、また民主主義社会において必要でない、と結論づけた。

(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。当該規定における禁止事項は、原則として、その日付の前または後に上市されまたはサービスが開始されたかにかかわらず、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、施行および制裁に関する章は、2025 年 8 月 2 日に適用開始となる。したがって、AI 法第 5 条の禁止事項に違反した場合の制裁に関する規定は、2025 年 8 月 2 日より前には適用されない。この暫定期間中は、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) それにもかかわらず、この暫定期間中においても、禁止事項は AI システムの提供者および導入者に完全に適用され、義務的となる。したがって、これらのオペレータは、AI 法第 5 条に基づき禁止される行為を構成し得る AI システムを上市し、サービスを開始し、使用しないように確保するために必要な措置を講じなければならない。モニタリングおよび制裁金に関する規定が後日まで適用されないとしても、禁止事項自体は直接的な効力を有し、したがって、影響を受ける当事者は国内の裁判所においてそれらを強制し、かつ禁止される行為に対し仮執行命令を求めることができる。

12. 欧州委員会のガイドラインの見直しおよび更新

(433) これらのガイドラインは、AI 法第 5 条における禁止事項の実例を交えた最初の解釈を構成する。欧州委員会は、禁止事項の理解のためにオペレータや当局に対し追加的支援を行い、かつ、AI システムの提供者および導入者、AI 委員会およびその他の関係する利害関係者からの意見を取り入れつつ更なる実務的なユースケースを継続的に収集する。

(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、ならびにこの分野における技術的、社会的、および規制上の発展のペースを考慮し、必要となり次第、これらのガイドラインを見直す。これには、市場監視の執行措置から得られるあらゆる関連する経験、ならびに禁止事項およびこれらのガイドラインにおいて検討された AI 法のその他の条項に関し欧州司法裁判所により示された解釈も含む。このような見直しにおいて、欧州委員会は、これらのガイドラインを撤回または修正することを決定することができる。欧州委員会は、AI システムの提供者および導入者、AI 委員会を通じて各国の市場監視当局、AI アドバイザリー・フォーラム、研究団体、市民社会団体に対し、将来の公聴の呼びかけに応じることにより、このプロセスに貢献するよう奨励する。

[244] AI 法第 111 条第 1 項および第 2 項参照。この条項は、適用除外条項が、AI 法第 113 条第 3 項(a)にいう AI 法第 5 条の適用を損なわないことを規定する。