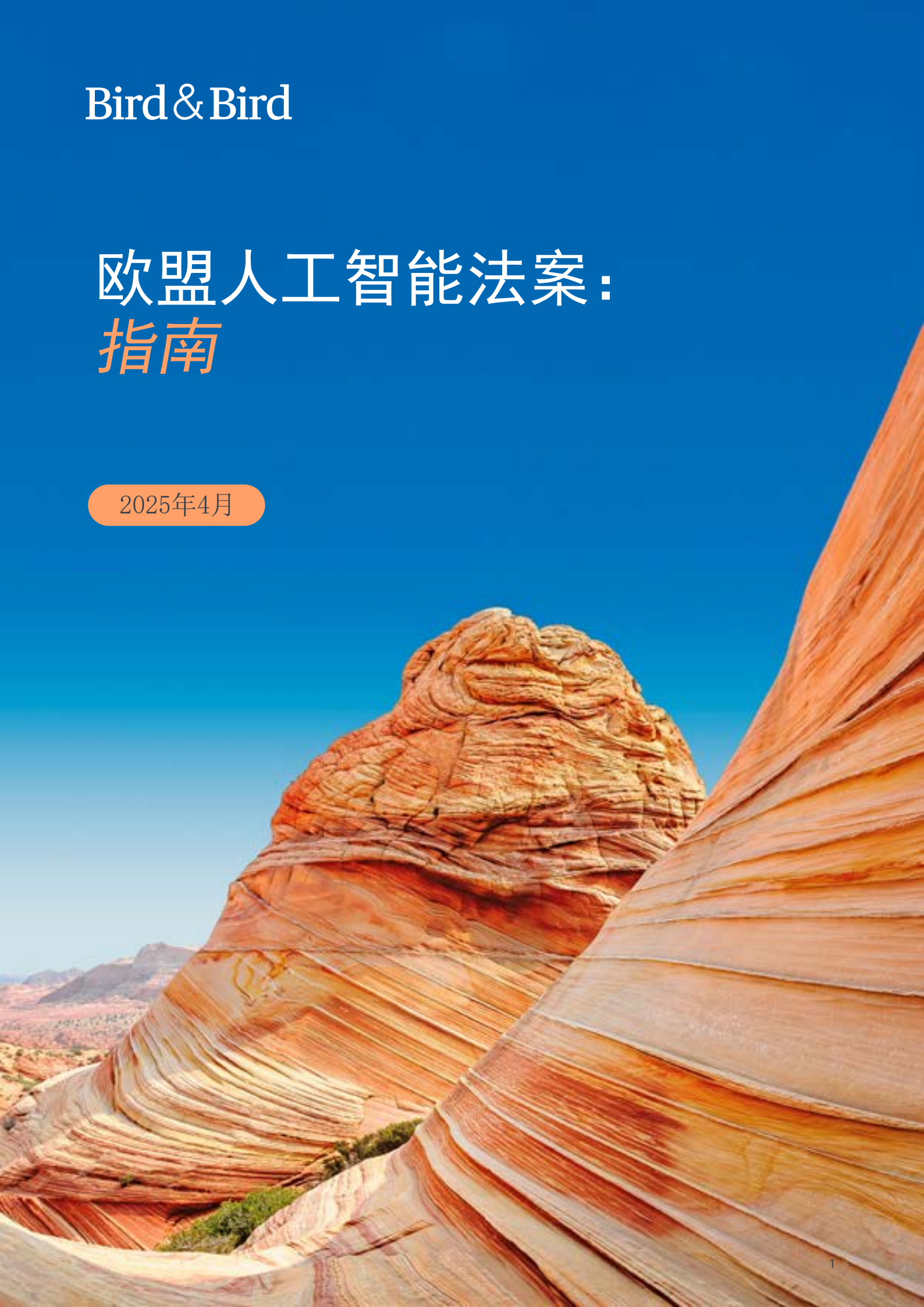
Bird&Bird

欧盟人工智能法案：*指南*

2025年4月



内容

概述、关键概念和实施时间

概述

欧盟（EU）是人工智能（AI）监管领域的先驱，其积极主动的举措为全球树立了标杆，确保人工智能发展合乎伦理和责任的标准。事实上，我们可能会见证一种新的布鲁塞尔效应（Brussels effect），类似于GDPR（《通用数据保护条例》）的影响力。欧盟的全面而审慎的框架优先强调透明度、问责制和人权。

《人工智能法案》（AI Act）的适用具有域外效力——无论提供者是否在欧盟内设立或位于第三国，许多规定都适用。《人工智能法案》适用于任何负责部署人工智能系统的提供者或实体，前提是“该系统产生的输出计划在欧盟使用”。外国供应商必须在欧盟内指定一名授权代表，以确保其遵守该法案的规定。但是，《人工智能法案》不适用于第三国的公共部门或与欧盟签订了警察和司法合作协议的国际组织，也不适用于为军事防御或国家安全目的投放市场的人工智能系统。这一广泛的范围旨在确保对人工智能系统及其使用进行全面监管。

您可以从本指南中得到什么

- 本章提供了整个《人工智能法案》的概述、其关键概念及其条款的适用日期。
- 第 2 章探讨了《人工智能法案》的地域范围和实质范围。
- 第 3、4、5 和 6 章讨论了《人工智能法案》对不同类型人工智能的要求——被禁止的实践做法；高风险系统；通用人工智能；以及需要更高透明度的人工智能。
- 第 7 章解释了《人工智能法案》在监管沙盒中测试人工智能的安排。
- 第 8 章讨论治理和执行。
- 第 9 章总结了《人工智能法案》通过后必须采取的众多进一步措施。
- 最后，第 10 章列出了本指南的所有贡献者。

基于风险的监管

欧盟对人工智能监管的方法以其基于风险的框架为特点。该法规采用技术中立的视角，根据风险等级对人工智能系统进行分类，从低风险到高风险不等。该体系确保高风险的人工智能应用（尤其是那些可能对基本权利产生重大影响的应用）被禁止或受到更严格的要求和监管。

欧盟高度重视推动负责任的人工智能的开发和使用。《人工智能法案》规定了严格的数据安全和用户隐私保护措施，确保人工智能系统的设计和部署以这些考虑为先。其中包括对如何处理和保护数据的严格要求，确保用户的个人信息安全。

此外，《人工智能法案》要求对人工智能系统进行全面的风险评估。这些评估有助于识别和减轻与人工智能技术相关的潜在风险，强化人工智能提供者之间的透明度和问责制。欧盟通过将这些评估设为强制性要求，确保人工智能开发者能够充分理解并应对其技术的潜在影响。

这种积极的做法旨在通过保护用户的权利和福祉来建立公众对人工智能技术的信任。通过优先考虑数据安全、隐私和风险管理，欧盟力求让公众相信人工智能可以安全且符合伦理地使用。这种对负责任发展的关注有助于促进更广泛的社会认同和对人工智能技术的整合，最终惠及整个社会。《人工智能法案》的制定不仅是为人工智能系统制定法律，也旨在建立相应的伦理框架，确保各组织在建立人工智能系统时考虑人工智能系统对人类、其他企业、环境和我们生活的许多其他方面的影响。

伦理准则是《人工智能法案》的核心

《人工智能法案》明确以欧盟委员会于 2019 年发布的《可信赖人工智能的伦理准则》（Ethical Guidelines on Trustworthy AI）为基础。尽管这些准则本身不具备约束力，但其诸多原则已被直接纳入《人工智能法案》。其中最典型的例证是，《人工智能法案》在众多条款中直接援引了《欧盟基本权利宪章》（Charter of Fundamental Rights of the European Union）所载的基本权利。例如，高风险人工智能系

统是指那些对欧盟人民的健康、安全和基本权利产生重大不利影响的系统。

在许多情况下，要正确适用《人工智能法案》，就必须对基本权利面临的风险进行分析，这既包括法律问题，也包括伦理问题。因此，可以说伦理问题已被嵌入到《人工智能法案》之中。

治理

欧盟采取去中心化监管模式，促进与各国当局的合作。《人工智能法案》规定，欧洲人工智能办公室（AI Office）为独立实体，作为欧盟人工智能专业知识的中央主管部门，在实施法律框架中发挥着关键作用。该办公室将鼓励开发可信赖人工智能系统并支持国际合作。欧洲人工智能委员会将由每个成员国的一名代表组成，欧洲数据保护监督员将作为观察员参与。

人工智能办公室旨在推动并促进良好实践准则的创建、审查与调整，同时考虑国际上的相关做法。为确保这些准则能够体现最新的技术发展水平并融合多元化的观点，该办公室将与相关的国家机构展开合作，并且可能会咨询民间社会组织、利益相关者

以及专家的意见，包括科学领域的专家。

关键概念

人工智能系统（另见第 2 章）

> *《人工智能法案》的大部分内容适用于“人工智能系统”，该法案将其定义为“一种设计用于以不同自主程度运行的基于机器的系统，该系统在部署后可能会表现出适应性，并且为了明确或隐含的目标，根据其接收到的输入推断如何生成输出，例如预测、内容、建议或决策，这些输出能够影响物理或虚拟环境”。*

值得注意的是，《人工智能法案》并未定义“人工智能”，而仅定义了“人工智能系统”一词。人工智能系统的定义有意与经济合作与发展组织（OECD）对人工智能系统的定义保持一致。该定义未提及任何特定技术或目前已知的人工智能系统方法。鉴于人工智能的快速发展特性，这可防止《人工智能法案》因技术发展而过时。

人工智能系统的“*推理*”能力应该是定义中的关键要素，这种能力可以明确区分人工智能系统和传统软件。如果一个计算机程序按照程序员预先定义的规则运行，它就不是人工智能系统；如果一个系统是使用允许程序基于提供给程序的输入数据或数据集自行创建规则的技术构建的，那么它就是人工智能系统。AI系统的定义在委员会于2025年2月6日发布的指南中被进一步讨论。

全供应链义务 （另见第二章）

人工智能法案适用于整个供应链中的所有参与者，从“*提供者*”开始，还包括系统的“*进口商*”、“*分销商*”和“*部署者*”。大多数责任由提供者承担，其次是部署者。

如果进口商、分销商或部署商在高风险人工智能系统上署上自己的名称或商标，他们就可能成为该系统的提供者。如果他们对人工智能系统进行重大修改或修改其预期用途，使其成为高风险系统，他们就可能成为高风险系统的提供者（见第 5 页）。

人工智能系统分类的风险方法

《人工智能法案》将风险定义为“*损害发生的可能性和损害的严重程度的结合*”。

基于风险的人工智能系统分类是《人工智能法案》的一个基本方面，重点关注人工智能系统可能对健康、安全和基本人权造成的潜在危害。这种方法将人工智能系统分为四个不同的风险等级：

1. 不可接受的风险：带来如此重大风险的人工智能系统是不可接受的，因此被禁止。
2. 高风险：高风险的人工智能系统受到严格的监管要求。
3. 有限风险：此类别的人工智能系统具有有限的风险，但有特定的透明度义务。
4. 风险极小或无风险：风险极小或无风险的人工智能系统不受《人工智能法案》的监管限制。

不可接受的风险：禁止的人工智能行为（另见第 3 章）

《人工智能法案》列出了禁止的人工智能行为，应理解为禁止将采用任何此类行为的人工智能系统投放市场、投入服务。该清单禁止：

- 使用潜意识技巧或故意操纵或欺骗性技术，严重扭曲行为，导致重大危害；

- 利用自然人或群体由于其特定特征而存在的弱点，造成重大危害；

- 社会评分系统，即根据自然人或群体的社会行为或个人特征对其进行评估或分类，从而导致不利或不公平的对待；

- 仅根据个人特征或个人资料来评估某人实施刑事犯罪的可能性；但用于支持基于与犯罪活动相关的客观和可验证的事实进行人为评估的情况除外；

- 基于从互联网或闭路电视（CCTV）无针对性采集信息面部识别数据库；

- 在工作场所或教育机构中做情绪推断，但出于医疗或安全原因的情况除外；

- 生物特征分类系统根据敏感数据对人员进行分类，但不包括标记或过滤合法获取的生物特征数据集（例如执法领域的图像）；

- 在公共场所用于执法目的的实时远程生物特征识别系统，某些特定情况除外。

在某些情况下，《人工智能法案》包含例外条款，允许在某些情况下做出这些“被禁止”的行为。实时生物识别就是一个很好的例子，该法规允许在特殊情况下适用实时生物识别。适用这些例外条款需要通知或事先授权。委员会于2025年2月4日发布了关于禁止性AI实践的指南。

高风险人工智能系统（另见第 4 章）

对高风险人工智能系统的全面监管构成了《人工智能法案》的主要部分。如果人工智能系统对欧盟内人员的健康、安全和基本权利产生重大不利影响，则被认定为高风险人工智能系统。不同的高风险人工智能系统适用不同的监管方式，具体可分为两类：

- 拟用作欧盟统一立法所涵盖的产品或产品安全组件的人工智能系统，例如民用航空、车辆安全、船用设备、无线电设备、玩具、电梯、压力设备、医疗设备、个人防护设备（列于《人工智能法案》附件一中）。

- 附件 III 所列的人工智能系统清单，其中包括用于教育、就业、信用评分、执法、移民、远程生物特征识别系统的人工智能，以及用作关键基础设施安全组件的人工智能系统。该清单可由委员会修订。

第一类高风险系统同时受到统一立法和人工智能法案的约束。提供者可以选择将《人工智能法案》的要求整合到附件 I 第A部分所列的欧盟统一立法所要求的程序中。此外，只有《人工智能法案》中的部分条款适用于与附件I中B部分列出的欧盟统一立法涵盖的产品（如航空设备）相关的高风险人工智能系统。

委员会将不迟于 2026 年 2 月 2 日提供高风险人工智能系统分类方面的实际帮助，包括一份全面的高风险和非高风险人工智能系统案例的综合实例清单。

高风险人工智能系统的资格认定例外

如果附件 III 所列的高风险人工智能系统不会对自然人的健康、安全或基本权利造成重大损害风险，包括不会对决策结果产生实质性影响，则不会被视为高风险人工智能系统。

此类情况仅在人工智能系统旨在实现以下四种情况时才会出现：

- 执行狭义的程序性任务；

- 改进先前完成的人类活动的成果；

- 检测决策模式或与先前决策模式的偏差，而不是为了在未经适当人工审查的情况下取代或影响先前完成的人工评估；或

- 执行与附件三所列用例目的相关的评估的准备任务。

然而，如果人工智能系统对自然人进行特征分析，则它始终被视为高风险人工智能系统，不能属于上述例外情况之一。

这项豁免很可能在实践中发挥重要作用，因为它可以避免将高风险人工智能系统投放市场所产生的义务和成本。例如，一个选项是剔除可以利用这种豁免的人工智能系统部分，以限制高风险人工智能系统的范围。

然而，即使提供者依赖着这项豁免，其对系统的评估也必须记录在案，并且在系统投放市场或投入使用之前仍必须在欧盟高风险系统数据库中注册。

高风险人工智能系统的全面义务

高风险人工智能系统提供者必须满足严格的要求。这些要求特别包括需要记录人工智能系统开发的每个阶段，使用高质量数据进行训练，制作系统文档以向用

户提供有关系统性质和目的的完整信息，或确保系统的准确性、稳健性和网络安全。高风险人工智能系统还必须在欧盟数据库中注册，并向公众开放。

人工智能系统供应链的义务

《人工智能法案》规定了高风险系统整个生命周期内供应链所有参与者的义务。责任不仅是“提供者”的责任，也是系统的“进口商”、“经销商”和“部署者”的责任，尽管大多数责任由提供者和部署者承担。

进口商和经销商的主要职责是核实进口或分销的高风险人工智能系统是否符合《人工智能法案》的要求。此外，如果进口商、分销商或部署者在系统上署上自己的名称或商标，对系统进行重大修改，或修改了人工智能系统的预期用途，导致系统具有高风险，则他们就可能成为高风险人工智能系统的提供者。

通用人工智能模型（另见第 5 章）

区分人工智能模型和人工智能系统在《人工智能法案》的应用中至关重要。人工智能模型是人工智能系统的重要组成部分，但它们本身并不构成人工智能系统。人工智能模型需要添加其他组件（例如用户界面）才能成为人工智能系统。《人工智能法案》主要规范人工智能系统，而不是模型。但是，它确实包含有关通用人工智能模型的规则。

《人工智能法案》规定了针对所有通用人工智能模型的规则，以及针对构成系统性风险的通用人工智能模型的附加规则。这些规则适用于以下情况：

- 通用人工智能模型提供者将其自身模型集成到自身的人工智能系统中，并投放市场或者投入使用；
- 通用人工智能模型的提供者仅向人工智能系统提供者提供自己的模型。

如果一个提供者的通用人工智能模型被用于第二个提供者的通用人工智能系统，而该系统又被集成到由第三个提供者建立的另一个具有更具体目的的人工智能系统中，那么这种区分可能尤为重要。

透明度义务 （另见第六章）

《人工智能法案》规定了四类人工智能系统的透明度义务：

- 旨在与自然人直接交互的人工智能系统；
- 包括通用人工智能系统在内的生成合成音频、图像、视频或文本内容的人工智能系统；
- 情绪识别或生物识别分类系统；
- 生成或操纵深度伪造的图像、音频或视频的人工智能系统。

在所有这些情况下，必须告知用户人工智能系统的使用情况。还有更详细的义务，例如以机器可读的方式标记输出，以便识别其是人工生成还是操纵的。

复杂的监督和执行结构（另见第 8 章）

《人工智能法案》规定了一个复杂的、多层次的监督执行结构。它既包括国家级实体，也包括欧盟级实体。每个级别都会有几类的机构，例如通知机构和指定机构、合格评定机构、欧洲人工智能委员会、人工智能办公室、国家主管机构和市场监督机构。

这些机构不仅将控制合规性，还将通过制定行为准则、组织人工智能监管沙盒以及为中小企业和初创企业提供支持等方式为市场提供支持。

技术标准、行为准则和指南的作用（另见第 7、8 和 9 章）

《人工智能法案》要求高风险人工智能系统的提供商加贴欧洲合格(CE) 标志。CE 标志表明其符合《人工智能法案》的要求。要获得该标志，提供者必须采用一致的技术标准。此外，符合一致标准的高风险人工智能系统或通用人工智能模型应被推定为符合《人工智能法案》的要求，只要这些标准涵盖这些要求或义务。因此，《人工智能法案》中的一般性的规定将由技术标准补充，这些标准将提供符合《人工智能法案》的具体形式。因此，我们可以预期，CE 标志和技术标准将在《人工智能法案》的实际应用中发挥非常重要的作用。

行为准则也应发挥重要作用。如果市场参与者没有制定行为准则，委员会可以在实施法案中提供通用规则。委员会还可以通过实施法案批准行为准则，并使其在欧盟范围内具有普遍效力。此外，委员会还有义务制定有关该条例实际执行的若干指导方针。

因此，《人工智能法案》可以被视为一个框架，更详细的义务将通过许多进一步的文件和法案规制。

执行（另见第 8 章）

《人工智能法案》规定了对不遵守规定的行为的严厉处罚，具体处罚取决于违规性质和所涉实体的规模。可能招致高额处罚的行为包括：

- 不遵守第 5 条中有关禁止人工智能实践的规定。在这种情况下，违法者可能面临最高 35,000,000 欧元或最高全球年营业额7%的罚款（以较高者为准）。
- 与数据、数据治理和透明度相关的违规行为：人工智能系统如被发现违反这些规定，最高可被罚款 2000 万欧元或全球年营业额的 4%。
- 不遵守第 99 条规定的任何条款（如与高风险人工智能系统有关的条款），将被处以最高达 1500 万欧元的罚款，如果违法者是一家公司，则最高可达其上一财年全球营业额的 3%，以较高者为准。

这些处罚凸显了遵守《人工智能法案》规定的重要性。企业必须充分了解这些处罚措施，并确保其人工智能系统符合该法案的要求。

时间线

《人工智能法案》将分阶段适用。对于在特定日期之前投放市场或投入使用的人工智能系统，该法案还设有过渡性安排。《人工智能法案》适用于 2026 年 8 月 2 日之前投放市场或投入服务的所有高风险人工智能系统的运营者，除非这些系统随后的设计发生重大变化（这种情况下，下列规定将全面适用于重新设计后的系统）。具体适用日期如下。

2024 年 7 月 12 日	《人工智能法案》在欧盟官方公报上公布，确定了法规中具体条款的适用日期。
2025 年 2 月 2 日	禁止实践禁令（第二章）。 人工智能素养规则（第 4 条）。
2025 年 5 月 2 日	通用人工智能的实践准则必须上线（第 56(9)条）。
2025 年 8 月 2 日	指定的国家主管部门（第三章第四节）。 通用人工智能（GPAI）的义务（第五章）。 治理（在欧盟和国家层面）（第七章）。 保密和处罚（与通用人工智能无关）（第十二章）。
2026 年 8 月 2 日	《人工智能法案》的所有其他规定适用（除非下文适用更晚的日期）。
2027 年 8 月 2 日	附件 I 所列的高风险类别。 2025年8月2日之前投放市场的通用人工智能模型（第111条）。
2030 年 8 月 2 日	高风险人工智能系统（下文所列系统除外），已于2026年8月2日之前投放市场或投入服务，且旨在供公共部门使用的（第111条）。
2030 年 12 月 31 日	附件 X 所列的大型信息系统组件，于2027年8月2日之前投放市场或投入服务（第111条）。

实质范围和地域范围

概览

- 《人工智能法案》涵盖人工智能系统、通用人工智能模型和禁止的人工智能实践。
- 可以对六类经济运营者（operator）施加义务：提供者（provider）、进口商（importer）、经销商（distributor）、产品制造商（product manufacturer）、授权代表（authorised representative）和部署者（deployer）。
- 涉及高风险人工智能系统的经济运营者负有重要义务。特定类别的人工智能系统的提供者和部署者也须履行透明度义务。
- 通用人工智能模型的提供者须承担义务。
- 《人工智能法案》适用于将人工智能系统或通用人工智能模型投放于欧盟市场、在欧盟投入服务、进口到欧盟或在欧盟经销的情形。该法案还适用于在欧盟设有机构或位于欧盟的部署者使用人工智能系统的情形。
- 属于《人工智能法案》适用范围内的人工智能系统提供者和部署者自 2025 年 2 月 2 日起须遵守人工智能素养（AI literacy）要求。

行动指南

- [ ] 请确认您、您的供应商或客户是否属于《人工智能法案》适用范围内的经济运营者（包括实质范围和地域范围）。
- [ ] 如果您或您的供应链上下游主体属于《人工智能法案》的适用范围，请检查是否存在任何人工智能系统或模型属于受监管的类别之一。
- [ ] 如果您是《人工智能法案》适用范围内的人工智能系统的提供者或部署者，请确保您已采取措施遵守该法案的人工智能素养要求。

实质范围

《人工智能法案》主要针对人工智能系统的“投放市场（placing on the market）”、“投入服务（putting into service）”和“使用（use）”制定统一规则。该法案对“高风险”人工智能系统施加了广泛的义务，并对某些人工智能系统规定了透明度义务。该法案还禁止某些人工智能实践，并对欧盟通用人工智能模型的供应进行监管。

《人工智能法案》还规定了市场监测、监督、治理和执法的规则，包括行政罚款以及支持创新的措施，尤其关注中小企业，例如通过人工智能沙盒的运作促进中小企业的发展。该法案还设立了两个新机构：(i) 欧洲人工智能委员会，负责向欧盟委员会和欧盟成员国提供建议并协助法案实施，以促进《人工智能法案》的一致性和有效性；(ii) 人工智能办公室，隶属于欧盟委员会，负责落实《人工智能法案》，促进可信赖的人工智能的开发和利用，并推动国际合作。

大多数运营者的定义参考了《人工智能法案》附件I中引用的欧盟产品立法中的三个关键术语：“在市场上提供（making available）”、“投放市场”和“投入服务”。

> *“在市场上提供”*是指在商业活动中，无论是否收费，将人工智能系统或通用人工智能模型提供给欧盟市场进行经销或使用；
>
> *“投放市场”*是指在欧盟市场上首次提供人工智能系统或通用人工智能模型；
>
> *“投入服务”*是指将人工智能系统直接提供给部署者首次使用，或供其在欧盟内按预期目的自用。

监管对象：运营者

《人工智能法案》规定了六类实体的义务：提供者、部署者、进口商、经销商、产品制造商和授权代表——这些实体统称为“运营者”。任何人工智能系统或通用人工智能模型都至少会有一个提供者。是否涉及其他运营者，取决于人工智能系统或通用人工智能模型的供应和部署方式。

《人工智能法案》未定义*“使用”*一词。从本质上讲，*“使用”*可理解为人工智能系统的核心特征，即从接收到的输入中推断并生成输出。这三个术语在委员会关于禁止人工智能行为的指南第2.3节中进行了讨论，该部分提供了每种行为的示例，以说明禁止行为的相关限制。

《人工智能法案》监管的运营者包括：

运营者		角色
适用于人工智能系统和通用人工智能模型	提供者（第 3(3) 条）	开发人工智能系统或通用人工智能模型，或拥有已开发人工智能系统或通用人工智能模型，并将其投放市场或以自己的名义或商标投入服务，无论是否收费。 尽管“投放市场”定义中提到的是欧盟市场，但如果主体没有将人工智能系统投放于欧盟市场，仅在欧盟内使用该系统的输出，该主体仍可被视为受《人工智能法案》监管的提供者。请参阅下文*“地域范围”*。 提供者可以是自然人或法人、公共机关、机构或其他组织。欧盟机构、组织、办公室和代理机构也可以作为人工智能系统的提供者。 如果某人工智能系统已由其他提供者投放市场或投入服务，主体仍然可以通过采取第 25(1)(a)-(c) 条规定的步骤之一成为提供者。请参阅下文*“高风险人工智能系统”*部分。
	授权代表（第 3(5) 条）	由在欧盟境外设立的提供者指定的设立于欧盟境内的自然人或法人，作为提供者的授权代表行事。其职责包括确保《人工智能法案》要求的文件可提供给主管当局，并与这些当局合作。请参阅第 22 条（适用于高风险人工智能系统）和第 54 条（适用于通用人工智能模型）。

仅适用于人工智能系统	部署者（第 3(4) 条）	在其权限下使用人工智能系统（不包括在个人非职业活动中使用）的任何自然人或法人、公共机关、机构或其他组织。欧盟机构、组织、办公室和代理机构也可以作为人工智能系统的部署者。
	进口商（第 3(6) 条）	位于欧盟或在欧盟境内设立的自然人或法人，将带有非欧盟设立者名称或商标的人工智能系统投放于欧盟市场。
	经销商（第 3(7) 条）	供应链中除提供者或进口商以外，向欧盟市场提供人工智能系统的自然人或法人。
	产品制造商（第 25(3) 条）	在某些情况下，如果符合以下条件，产品制造商将被视为高风险人工智能系统的*“提供者”*： 该人工智能系统是《人工智能法案》所涵盖的产品的安全组件（根据附件 I 第 A 部分中引用的欧盟产品安全立法）；并且 制造商将该人工智能系统与该产品一起以自身名义或商标投放于欧盟市场或投入服务。 《人工智能法案》中没有对*“产品制造商”*一词进行定义，但序言第 87 条中明确指出，这一术语指的是根据《人工智能法案》附件I中引用的欧盟产品安全立法中定义的“制造商”。

《人工智能法案》下的间接义务

《人工智能法案》对高风险人工智能系统提供者的组件供应商施加了间接义务。向高风险人工智能系统供应人工智能系统、工具、服务、组件或流程的供应商，必须与高风险人工智能系统的提供者签订书面协议，并帮助后者遵守《人工智能法案》下的义务（第 25(4) 条）。这项义务不适用于根据免费和开源许可向公众提供此类工具、服务、流程或组件（通用人工智能模型除外）的第三方。

《人工智能法案》赋予的权利

与为个人提供全面权利的《通用数据保护条例》（GDPR）不同，《人工智能法案》赋予个人的权利是有限的。《人工智能法案》仅赋予欧盟境内的受影响人员关于个体决策的解释权（第 86 条）。受影响人员是指那些受到基于附件 III 中确定的高风险人工智能系统输出的决策影响的人员，这些决策对其具有法律效力或类似的重大影响。此处使用的措辞与 GDPR 的自动化决策条款（GDPR 第 22 条）中使用的措辞相似，但这两项条款的适用范围并不完全相同。

监管主题：人工智能系统

《人工智能系统》第 3(1) 条对人工智能系统进行了广泛的定义，即：*“一种基于机器的系统，旨在以不同程度的自主性运行，并可能在部署后表现出适应性的机器系统，且为了明确或隐含的目标，根据其接收的输入中推断如何生成可以影响物理或虚拟环境的输出，如预测、内容、建议或决策”*。

此定义旨在与OECD人工智能原则（OECD AI Principles）中的定义保持一致。人工智能系统的一个关键特征是其推断能力，即从输入或数据中获取输出并推导出模型和/或算法。而传统的软件，仅根据自然人定义的规则执行操作，本身并不被视为人工智能系统。人工智能系统可以独立使用，也可以作为产品的组件使用，无论该人工智能系统是物理集成到产品中还是以非集成的方式服务于产品的功能。

2025年2月，委员会发布了关于该定义的指南。该指南对定义的各个方面进行了进一步解释，并特别强调了“推理能力”。在正面阐述方面，指南概述了各种实现该能力的机器学习方法。同时，指南列举了一些不具备该能力的系统，特别是主要基于数学或统计方法的系统，并指出它们不应纳入《人工智能法案》的适用范围。其中一个值得注意的负面示例是“逻辑回归”，该方法在金融领域被广泛应用。

根据《人工智能法案》，人工智能系统可以分为以下类别：

- 高风险人工智能系统；
- 具有透明度风险的人工智能系统（有限风险人工智能系统）；以及
- 其他所有人工智能系统。

人工智能系统也可能成为禁止的人工智能实践的一部分。这可能是由于该人工智能系统的某些功能，也可能是因为该人工智能系统的使用方式。

高风险人工智能系统

《人工智能法案》第III部分规定了高风险人工智能系统。这些人工智能系统对欧盟境内人员的健康、安全和基本权利构成重大损害风险。人工智能系统可通过以下两种方式被归类为高风险系统：

- 第 6(1) 条：人工智能系统作为产品的安全组件使用，该产品受到某些欧盟产品安全立法（《人工智能法案》附件I所列的欧盟统一立法）的监管并须根据此类立法与第三方评估机构进行符合性评估程序，或该人工智能系统自行构成此类产品（例如，用于医疗诊断目的的人工智能系统本身将构成受监管的医疗器械）；或
- 第6(2)条：人工智能系统属于《人工智能法案》附件III所列的八个类别之一——除非提供者能够证明并记录该人工智能系统不会构成重大损害风险。

大多数与高风险人工智能系统相关的义务由提供者（包括前述所提到的产品制造商）承担，而部署者、进口商和经销商以及相关的授权代表承担的义务则相对有限。

请参阅本指南第 4 章了解更多详细信息。

有限风险人工智能系统

《人工智能法案》对以下角色规定了某些透明度义务：

- 旨在与自然人直接互动的人工智能系统的提供者（第50(1)条）；
- 生成合成音频、图像、视频或文本内容的人工智能系统的提供者（第50(2)条）；
- 情绪识别系统或生物特征分类系统的部署者（第50(3)条）；
- 生成或操纵构成深度伪造（deep fake）的图像、音频或视频内容的人工智能系统的部署者（第50(4)条）。

请参阅本指南的第 6 章了解更多详细信息。

其他所有人工智能系统

所有不属于上述类别并且未用于禁止的人工智能实践的人工智能系统，不受《人工智能法案》规定的直接法律义务的约束。未来可能会制定涵盖这一更广泛类别的人工智能系统及其部署者的自愿行为准则（第 95 条）。提供者和部署者可以选择遵守这些行为准则。

除了与特定类别的人工智能系统相关的规定外，符合《人工智能法案》规定的人工智能系统的提供者或部署者，还必须采取人工智能素养措施，确保其员工和代表其操作和使用人工智能系统的其他人员拥有足够的知识、技能和理解，能够了解人工智能系统的部署、机会和风险（第 4 条）。这项义务旨在促进欧盟以可信赖的方式开发、运营和使用人工智能——不过需要注意的是，这项条款涉及自愿行为准则，不遵守人工智能素养义务不会受到行政罚款。

监管主题：禁止的人工智能实践

《人工智能法案》将具有某些禁止功能和/或旨在用于某些禁止目的的人工智能系统的投放市场、投入服务和使用行为规定为禁止的人工智能实践，例如通过从互联网或闭路电视（CCTV）中无差别地抓取面部图像来创建或扩展面部识别数据库的人工智能系统。这些做法被认为极其有害和滥用，违背了欧盟价值观和基本权利。禁止的人工智能实践列在《人工智能法案》第 5 条中。此列表不妨碍其他欧盟法律（如数据保护、禁止歧视、消费者保护和竞争法）中关于人工智能实践的禁止性规定。

请参阅本指南第 3 章了解更多详细信息。

监管主题：通用人工智能模型

> 《人工智能法案》第3(63)条将通用人工智能模型定义为：*“包括使用大量数据进行大规模自我监督训练的人工智能模型，无论该模型如何投放市场，都具有显著的通用性，能够胜任各种不同的任务，并且可以集成到各种下游系统或应用中，但在投放市场前用于研究、开发和原型设计的人工智能模型除外”*。

《人工智能法案》并未对“人工智能模型”作出定义。序言第 97 条指出，尽管人工智能模型是人工智能系统的基本组件，但它们本身并不构成人工智能系统，需要用户界面等其他组件才能成为人工智能系统。序言第 98 条和第 99 条进一步讨论了通用人工智能模型的特征。

《人工智能法案》对通用人工智能模型进行监管，并对具有系统性风险的通用人工智能模型规定了额外义务。这些规定适用于被投放市场的通用人工智能模型的提供者：投放市场可以通过多种方式实现，例如通过库、API、直接下载或物理副本。

序言第 97 条指出关于通用人工智能模型的规定同样适用于当这些模型被集成到人工智能系统中或成为人工智能系统的一部分时的情形。当通用人工智能模型的提供者将自己的模型集成到自己的人工智能系统中，并将该系统在市场上提供或投入服务时，序言第 97 条指出该模型将被视为已投放市场。此时，除有关人工智能系统的条款外，通用人工智能模型的条款也将适用。将第三方通用人工智能模型集成到自己的人工智能系统中的人被视为“下游提供者（downstream providers）”，并在《人工智能法案》下享有某些权利。然而，《人工智能法案》似乎设想，对第三方的通用人工智能模型进行微调并将该微调后的模型集成到自己的人工智能系统中（或以其他方式将微调后的通用人工智能模型投放市场或投入服务）的提供者，仅就该微调而言，将被视为该模型的提供者（参见序言第 109条）。

请参阅本指南第 5 章了解更多详细信息。

地域范围

人工智能系统相关条款

《人工智能法案》为其人工智能系统相关条款设定了广泛的司法管辖范围：当人工智能系统（无论是独立存在，还是作为《人工智能法案》附件I中欧盟产品安全立法所涵盖的产品的组件存在）符合以下任一情况时，这些条款将适用：

- 被投放于欧盟市场、在欧盟投入服务、进口到欧盟或在欧盟经销；或
- 由在欧盟设有机构或位于欧盟的部署者使用。

第一点在适用时将不考虑人工智能系统的提供者设立于何处。《人工智能法案》并未定义“设立（establishment）”的概念，其解释预计将与其他欧盟立法（例如 GDPR）中类似，即范围较为广泛。

除上述情况外，如果欧盟以外的人工智能系统产生的输出在欧盟境内被使用，则人工智能系统相关条款也将适用。

此种情况下，非在欧盟设立/位于欧盟的提供者和部署者也将受到《人工智能法案》的约束。序言第 22 条进一步明确，即使相关人工智能系统未在欧盟投放市场、投入服务或使用，《人工智能法案》在此类情况下仍适用。

禁止的人工智能实践

《人工智能法案》关于禁止的人工智能实践的条款适用于第5条中规定的相关人工智能实践投放于欧盟市场、在欧盟投入服务以及使用的行为。如前所述，*“投放市场”*和*“投入服务”*的定义均针对欧盟市场。《人工智能法案》并未具体说明禁止的*“使用”*包含哪些内容，例如是否包括在欧盟境内仅使用人工智能系统输出的情形。委员会关于禁止性人工智能实践的指南建议，*“使用”*应以广义理解，以涵盖在人工智能系统投放市场或投入使用后的任何生命周期阶段的使用或部署。此外，指南还指出，*“使用”*还可能包括将人工智能系统集成到使用该系统的个人或实体的服务和流程中，包括作为更复杂的系统、流程或基础设施的一部分。

通用人工智能模型

《人工智能法案》的通用人工智能模型相关条款适用于通用人工智能模型提供者将其投放于欧盟市场或在欧盟投入服务的情形，无论该提供者位于/设立于何处。

法条指引

实质范围	第 1 条	序言 1-3，6-8
实质范围	第 2 条	序言 9-11

适用例外

某些活动完全超出了《人工智能法案》的适用范围，包括以下情况：

- 欧盟法律范围之外的领域（如涉及国家安全的活动）。无论成员国委托何种实体执行与这些豁免活动相关的任务，均不适用该法案。鉴于《欧盟运作条约》规定了欧盟的广泛权限，实际上该条款的适用范围非常有限；

- 专门用于军事、国防和国家安全目的的人工智能系统，无论是否经过修改以及从事这些活动的实体类型为何，将其在欧盟投放市场、投入服务、使用或仅在欧盟境内使用其输出的情形均不适用《人工智能法案》。如果投放市场或投入服务的人工智能系统既用于例外用途（军事、国防或国家安全），又用于非例外用途（如民用或执法目的），则仍需遵守《人工智能法案》，并由系统的提供者确保合规；

- 第三国公共机关或国际组织在与欧盟或欧盟成员国进行执法和司法合作的国际合作或协议框架内使用人工智能系统的情形，但前提是该第三国或国际组织为保护个人的基本权利和自由提供了充分的保障。使用人工智能系统输出的国家当局和欧盟机构、组织、办公室和代理机构仍需遵守欧盟法律；

- 专为科学研发目的而开发并投入服务的人工智能系统和模型，包括其输出；

- 人工智能系统或模型在投放市场或投入服务之前的研究、测试或开发，但在真实环境中的测试除外；

- 个人部署者在纯粹的个人非职业活动中使用人工智能系统。这与 GDPR 的“家庭豁免”类似，但这些人工智能系统的提供者仍需遵守《人工智能法案》；

- 在免费和开源许可下发布的人工智能系统，除非这些系统投放市场或投入服务时属于高风险人工智能系统、禁止的人工智能系统或被透明度义务涵盖的系统。

与其他监管框架的关系

- 作为一项条例，《人工智能法案》无需成员国转化立法，可直接适用于欧盟成员国。除非《人工智能法案》明确授权，否则欧盟成员国不得对人工智能系统的开发、营销和使用施加限制。这种情况仅限于少数情形，例如：欧盟成员国可针对远程生物特征识别系统（部分此类系统被列为禁止的人工智能实践（第 5(5) 条）以及事后远程生物特征识别系统（此类系统构成高风险的人工智能系统（第 26(10) 条））的使用制定更严格的法律。

- 《人工智能法案》关于高风险人工智能系统的条款是围绕欧盟产品的新立法框架（New Legislative Framework for EU products）制定的。该框架是一套关于产品投放于欧盟市场的规则，旨在强化市场监督、符合性评估及 CE 标志的规定，同时它还以工具箱的形式为工业产品建立了一个共同的法律框架，供未来立法使用。《人工智能法案》明确了这些新立法框架工具在人工智能系统中的具体适用方式。

- 同时，《人工智能法案》补充了欧盟统一立法——此为欧盟产品安全立法的集合。根据这些立法，某些人工智能系统将被归类为高风险。

- 《人工智能法案》的义务是对GDPR、《电子隐私指令》和《执法指令》规定义务的补充，且不影响后者的效力。

- 同时，《人工智能法案》补充了欧盟统一立法——此为欧盟产品安全立法的集合。根据这些立法，某些人工智能系统将被归类为高风险。

- 《人工智能法案》的义务是对GDPR、《电子隐私指令》和《执法指令》规定义务的补充，且不影响后者的效力。

禁止人工智能实践

概览

- 第五条列出了八种被认为具有不可接受的风险的应被禁止的实践。
- 禁令将于 2025 年 2 月 2 日生效。
- 被禁止的行为包括：
  - 潜意识、操纵性或欺骗性技术
  - 每起案件中利用弱势群体的技术，这些技术显著扭曲行为并带来重大伤害风险
  - 特定用例中的社交评分
  - 根据用户画像预测犯罪行为
  - 爬取互联网或闭路电视数据以获取面部识别数据库
  - 工作场所或学校的情绪推断
  - 通过生物特征分类推断种族、政治观点、工会成员身份、宗教或政治信仰、性生活或性取向
  - 用于执法目的的公共场所实时远程生物特征识别系统。
- 许多禁令都有例外——需要具体情况具体分析。
- 该清单并非最终版本：每年都会重新评估。
- 不合规行为将被处以最高达 3500 万欧元或上一财年全球年营业额总额的 7%（以较高者为准）的罚款。
- 这些禁令与运营者无关，无论参与者的角色如何（即提供者、部署者、分销商还是进口商）均适用。

待办事项

- [ ] 检查您使用的人工智能系统，看它们是否属于被禁止的类别。
- [ ] 每年检查此列表的更新，因为禁止实践的列表可能会随着时间推移而发生变化。
- [ ] 考虑是否有任何例外情况。被禁止的实践并非绝对；许多都有例外情况。

禁止人工智能实践

《人工智能法案》采用基于风险的监管方式，因此根据风险等级适用不同的要求。本章重点关注被禁止实践，即与欧盟价值观相冲突并明显威胁自由、平等和隐私等基本权利的行为。这些禁令是立法者为应对透明度和伦理关切以及确保人权保护所作出的努力。

被禁止实践在第 5 条中详尽列出（并在该法案的第 28 至 45 条以及委员会于2025年2月4日发布的指南中进一步解释），并为人工智能在欧盟内可以做什么和不可以做什么提供了明晰的框架。第 5 条中的禁令自 2025 年 2 月 2 日起生效，因此是首批生效的条款，突显了其重要性。

第 5 条中列出的禁止实践清单是详尽的，但并非最终版本。委员会每年将评估修订禁止实践清单的必要性（第 112 条），并可向欧洲议会和理事会提交评估结论。因此，禁止实践清单在未来可能会有所调整。

目前有八项禁止实践，重点关注那些严重扭曲人们行为或引起民主社会担忧的实践。生物特征识别系统受到特别关注。然而，许多禁令都有详细的例外情况，每项行为都应根据具体情况进行考虑。

第 5(1)(a) 条 潜意识、操纵性或欺骗性技术

第一项禁令涉及人工智能系统在下列情况下部署潜意识、操纵或欺骗技术：

- 这些技术的目的在于、或者实际上已经产生了严重扭曲自然人或群体行为的效果；
- 明显损害自然人或群体做出明智决定的能力；并且
- 导致他们做出原本不会做出的决定，并且导致或很可能导致他们遭受重大伤害。

序言 第29条 中明确提到的技术包括：部署潜意识组件，例如人们无法感知的音频、图像、视频刺激，或其他操纵或欺骗技术，以破坏或损害个人的自主权、决策权或自由选择能力的方式，使人们在不知不觉中受到这些技术的影响，或者即使人们意识到了，仍然可能被欺骗或无法控制或抵抗它们。序言第29条提到的具有实质性扭曲人类行为且造成严重有害影响能力的脑机接口，也可能是该法案试图规范使用神经数据的工具的一种尝试，目前在科罗拉多州、加利福尼亚州和智利等其他司法管辖区正在讨论此类数据的使用。

若要禁止一个人工智能系统，欺骗技术与造成的重大损害之间必须存在因果关系。在立法过程中增加了“重大”伤害的门槛，明确指出并非所有的欺骗性设计都会落入这一条款的适用范围。

该条款具有解释空间，尤其是“欺骗性”一词将引发进一步的讨论。根据委员会的指南，欺骗性技术可能包括提供虚假或误导性信息，以达到或产生误导个人的目的或效果，前提是符合第一项禁令的其他要求。

第 5(1)(b) 条 利用弱势群体

第二类禁止的人工智能实践旨在保护弱势群体。弱势群体包括三类：因年龄、残疾或特定社会或经济状况而处于弱势。

只有当人工智能系统的目的或效果是严重扭曲个人行为，并且以造成或可能造成重大伤害的方式进行时，才会被禁止。

根据委员会关于禁用行为的指南，从社会经济角度来看，如果某种情况可能发生在任何人身上，而不论其社会经济状况如何（例如，遭受不满或孤独），则不构成剥削。然而，在这种情况下，剥削行为可能受到《人工智能法案》第5(1)(a)条的约束。

由于训练数据的偏见，AI 系统可能会无意间影响社会经济弱势群体，但如果没有针对性的意图，则不会自动构成对弱势群体的剥削。然而，根据委员会关于禁用行为的指南，如果 AI 提供者或使用者明知其系统对社会经济弱势群体构成非法歧视，并且预见到严重危害而未采取纠正措施，他们仍可能被认定为利用了这些弱势群体的脆弱性。

如果利用人工智能系统寻找贫困人群，在经济上利用他们的弱点，则可能存在对个人经济状况的剥削。使用人工智能系统进行营销和销售的组织应确保根据这一要求完成系统测试。

“重大损害”这一概念既适用于潜意识技术，也适用于对弱势群体的利用。在立法过程中，曾要求损害必须是身体或心理上的这一要求已被取消。《人工智能法案》似乎试图采取一种宽泛的方法来定义伤害的概念，尽管第 29 条仍然给出了对身体和心理健康以及经济利益产生重大不利影响的例子。该条款还指出，伤害可以随着时间的推移而累积。

第 5(1)(c) 条 社会评分

第三项禁令涉及所谓的社会评分，即根据自然人或群体的社会行为或已知、推断或预测的个人或个性特征，在一定时期内对其进行评估或分类产生社会评分。在两种情况下禁止使用社会评分：

- 对特定自然人或群体产生有害或不利待遇，而这种待遇的社会场景与最初生成或收集数据的场景无关；以及
- 对特定自然人或群体产生有害或不利待遇，而这种待遇是不合理的，或与其社会行为或其严重性不成比例的。

全球多个政府都在使用社会评分。2021年，荷兰政府因风险评分算法存在缺陷而下台，导致政府被不合理地指控基于个人特征和行为对福利津贴实行欺诈。该案中的算法针对的是少数群体和基于经济状况被区分的群体。虽然政府可能是人们在考虑社会评分时首先想到的例子，但该规定更为广泛，涵盖了公共和私人领域的所有社会评分系统。许多算法本质上依赖于行为评分。然而，《人工智能法案》仅禁止那些在与社会无关的背景下导致不利待遇的评分系统。这一关键限制针对的是社会评分的后果，防止不公正的结果或对个人或特定群体的歧视。

因此，《人工智能法案》中的社会评分禁令取决于数据的获取背景和数据使用的背景。正如委员会关于禁止做法的指导文件所示，合法的活动，如金融服务中的信用和风险评分，如果能够提高服务质量或防止欺诈，是被允许的。相反，保险公司使用银行的消费和其他财务数据来设定人寿保险费用，则被视为非法的社会评分的例子。

第 5(1)(d) 条 犯罪风险评估的画像分析

第四项禁令是禁止将仅基于用户画像或评估某人的性格特质和特点的方式来判断或预测某人实施刑事犯罪可能性的人工智能系统投放市场、为此特定目的投入服务、或使用此类人工智能系统。用于支持评估某人是否参与犯罪活动的人工智能系统除外，此类评估基于与犯罪活动直接相关的客观和可验证的事实，即检测工具是基于事实的，是人类决策的补充，但不能取代人类决策。这项禁令旨在避免将尚未犯罪的人视为有罪的情况——正如电影《少数派报告》中所描述的那样。它与《基本权利宪章》第一条规定的人的尊严息息相关。

委员会关于禁止行为的指南强调，如果私人实体以公共权力行事或协助执法，它们也可能受到禁止。例如，如果满足特定标准，一家为执法机关分析数据的私人公司可能会面临禁止。

委员会的指南还建议，某些条件下，人工评估AI系统评估的回顾性审查可能不在适用范围内。这一点受欧盟法院（CJEU）判例法的影响，判例法强调人工审查的重要性，以确保基于客观标准和非歧视性原则的AI驱动决策，从而超出了《人工智能法案》初步豁免的范围。

第 5(1)(e) 条 人脸识别数据库

第五项禁止的实践是禁止将通过从互联网或闭路电视录像中无区别地爬取面部图像创建或扩展面部识别数据库的人工智能系统投放市场、为此特定目的投入服务、或使用此类人工智能系统。序言第 43 条认为这种做法强化了被大范围监视的感受，并可能导致对基本权利的严重侵犯，包括隐私权。这可能是对监管机构对 Clearview AI 进行调查的回应。

委员会关于禁止行为的指南澄清了关于面部识别数据库的几个关键点。这些数据库可以是临时的、集中式的或去中心化的，如果它们可以用于面部识别，无论其主要用途是什么，都属于《人工智能法案》第5条第1款第(e)项的范畴。目标化抓取，如收集特定个体的图像或使用反向图像搜索，是允许的，但与无目标抓取相结合的做法是被禁止的。该禁止令不涵盖对其他生物识别数据（如语音样本）的无目标抓取，也不包括未用于识别的数据库，例如用于AI模型训练且不涉及个人身份的数据库。

第 5(1)(f) 条 工作生活和教育中的情绪推断

第六项禁止实践是禁止将在工作场所和教育机构领域推断自然人的情绪的人工智能系统投放市场、为此特定目的投入服务或使用此类人工智能系统。但出于安全或医疗原因（例如用于治疗用途的系统）的情况除外。指南明确指出，学校和工作场所的定义应广泛解读，在工作场所使用的情况下，也应涵盖招聘的选择和录用阶段。另一方面，出于安全或医疗原因的例外应狭义解释。例如，用于衡量工作场所倦怠或抑郁的系统将不被豁免。

序言第 18 条对快乐、悲伤、愤怒等情绪或意图作了区分。它解释说，这个概念不包括身体状态，例如疼痛或疲劳（因此，用于检测专业飞行员或驾驶员疲劳状态以防止事故发生的系统不会受到影响）。它也不包括对皱眉或微笑等明显表情的检测，或对手部、手臂或头部动作等手势的检测，或对一个人声音特征的检测，例如提高声音或低语。然而，指南仍未明确“意图”的含义，这也是情感识别系统定义中涵盖的内容。

《人工智能法案》对*“情绪识别系统”*一词进行了明确定义，即*“基于生物特征数据识别或推断自然人的情绪或意图的人工智能系统”*。

奇怪的是，第 5(1)(f) 条没有使用这个术语，而使用的是以推断情绪为目标的人工智能系统（即没有要求必须从生物识别数据中得出）。然而，委员会的指南明确指出，第5(1)(f)条应被解读为指情感识别系统，这是根据该法案定义的术语。。他们进一步澄清，非生物特征情绪识别系统（例如基于文本的）在不与按键分析等生物特征数据结合使用的情况下不被禁止。该法案提到了生物特征情绪识别系统的不准确性及其在权力不平衡的环境中（如工作场所和学校）的侵入性作为其在这些环境中禁止使用的原因。然而，《人工智能法案》并没有解释为什么非生物特征情绪识别系统比生物特征系统侵入性更小或更准确。

第 5(1)(g) 条 生物特征识别分类

第七项禁令针对的是使用生物识别分类系统，该系统基于个人的生物识别数据对个体进行分类，以推断或推算GDPR下某些（而非全部）特殊类别数据，具体包括：种族、政治观点、工会会员身份、宗教或政治信仰、性生活或性取向。

GDPR 下的特殊类别数据不受该禁令的约束，包括种族来源、健康和遗传数据的推断。但是，推断这类数据可能会属于附件 III 中的高风险类别。此外，禁令不包括对合法获取的生物识别数据集进行标记或筛选，也不包括执法部门对生物识别数据进行分类（例如，执法部门根据发色或眼色对图像进行排序以搜寻嫌疑人）。然而，目前尚不清楚，使用生物识别分类进行此类标记和筛选的系统是否属于高风险类别，因为序言第54条暗示，旨在根据GDPR下的敏感属性或特殊类别数据进行生物识别分类的人工智能系统（只要这些系统未被《人工智能法案》禁止），应被归类为高风险，并且指南还指出，大多数符合《人工智能法案》第5条所列禁令例外的人工智能系统将被归类为高风险。这意味着被豁免的标签和过滤系统可能会被归入高风险类别。

序言第 16 条明确指出，生物识别分类系统不包括与另一项商业服务相关的纯辅助功能，这些功能由于客观技术原因，无法在没有主要服务的情况下使用，并且这不是为了规避《人工智能法案》规则的规避机制（例如，零售“先试后买”过滤器或社交媒体过滤器）。

指南还明确指出，生物特征分类的范围不包括根据服装或配饰（如围巾或十字架）或社交媒体活动进行的分类。

第 5(1)(h) 条 公共场所的实时远程生物特征识别

第八项也是最后一项禁令是禁止在公共场所使用实时远程生物特征识别系统（ “RBI” ）用于执法目的。RBI 系统是一种用于在未经个人参与的情况下，通常在一定距离外，通过将生物识别数据与参考数据库中的数据进行比对，以识别自然人的人工智能系统。人工智能法案未定义多少时间构成“显著延迟”。然而，指南建议，当个人可能已经离开采集生物特征数据的地点，并且无法让执法部门迅速作出反应时，通常可视为“显著延迟”的情况。

用于验证（即确认某人是其声称的人，访问服务、设备或安全进入场所）的生物特征识别系统与 RBI 不同，因此不受此禁令的约束（第 15 条）。指南明确指出，识别与验证之间的区别在于个人在过程中的主动参与程度，这可能对自然人的基本权利产生较小影响。然而，仅仅告知个人摄像头的存在并不足以构成主动参与，他们需要主动且有意识地走到特定安装的位置，使其能够积极参与验证过程。

《人工智能法案》允许（但不要求）成员国在有限的情况下使用 RBI 用于执法目的，但前提是使用 RBI 是以下严格必要的情形：

- 针对特定被绑架、人口贩卖或性剥削受害者的有针对性搜索，以及寻找失踪人员；
- 防止对生命或人身安全的特定、重大和迫切的威胁，或防止真实且当前或可预见的恐怖袭击威胁
- 定位或识别涉嫌刑事犯罪的人员，进行刑事调查、起诉或对严重罪行（即附件II中所指的罪行，且在相关成员国可被判处至少四年监禁）执行刑事处罚。

豁免仅允许使用 RBI 来确认特定目标个人的身份。此外，使用 RBI 应考虑所涉及情形的性质，特别是如果不使用该系统将造成的损害的严重性、概率和规模，以及使用该系统对有关人员的权利和自由的影响。进一步的保护措施包括需要完成基本权利评估、根据第 49 条在欧盟数据库中注册该系统，以及司法或行政机关对每个使用案例的事先授权（紧急措施除外）。此外，每次在公共场所使用 RBI 都必须通知相关市场监督机构和国家数据保护机构。国家机构必须向欧盟委员会报告，欧盟委员会则根据这些规定编制一份关于 RBI 使用情况的年度国家报告。

禁令适用主体

如第二章所述，《人工智能法案》对参与人工智能系统的不同主体进行了区分，根据他们在人工智能模型或系统中的角色赋予他们特定的责任。这种方法确保那些对人工智能技术开发和实施最具影响力的人遵守最高标准。

然而，禁止实践的规定与运营者无关。换句话说，这些规定具有普遍适用性，不依赖于主体的具体角色（即无论他们是否参与提供、开发、部署、分发或使用从事禁止行为的人工智能系统）。

这种广泛的适用范围凸显了该法案致力于制止可能侵犯基本权利或带来不可接受风险的实践的决心，强调了对所有有害人工智能技术类型的交互采取全面监管。

执行与罚款

当某种实践被禁止时，相关人工智能系统不得在欧盟使用。如果发生违法行为，主管部门可处以最高相当于违法者上一财年全球年营业额 7% 或 3500 万欧元的罚款，以较高者为准。

国家市场监督机构将负责确保遵守《人工智能法案》关于禁止人工智能系统的规定。他们将每年向欧盟委员会报告当年发生的禁止实践的情况以及他们采取的措施。

对应的法律规定条款

潜意识、操纵性或欺骗性技术	第5(1)(a)条	序言 28 & 29
利用弱势群体	第5(1)(b)条	序言 28 & 29
社交评分	第5(1)(c)条	序言 31
犯罪风险评估的画像分析	第5(1)(d)条	序言42
人脸识别数据库	第5(1)(e)条	序言43
工作生活和教育中的情绪推断	第5(1)(f)条	序言 44 - 45
生物特征识别分类	第5(1)(g)条	序言 30
公共场所实时远程生物特征识别	第5(1)(h)条	序言 32 - 41

其他相关参考来源

- 《委员会关于《人工智能法案》（欧盟法规 2024/1689）所规定的被禁止 人工智能实践的指南》
- 可信赖人工智能的伦理准则：人工智能高级专家组 (2019)
- EDPB 关于通过视频设备处理个人数据的指南
- EDPB关于在执法领域使用面部识别技术的指南
- EDPB 关于自动决策和画像分析的指南
- EDPB关于人工智能法案提案的联合意见
- EDPB关于社交媒体欺骗性设计模式的指南
- 芬兰市场管理局关于欺骗性设计的指南

高风险人工智能系统

概览

- 如果人工智能系统旨在用于以下用途，则属于“高风险”系统：

  — 作为产品本身或者产品的安全部件，并且根据附件I所涵盖的立法必须进行第三方符合性评估；或者

  — 用于附件III中描述的目的之一。

- 根据《人工智能法案》，高风险人工智能系统的提供者、部署者、进口商、经销商以及供应商在提供产品或服务的过程中均需承担相应义务。市场主体可以同时担任多个角色，但也需同时遵守不同角色项下的合规义务。

- 高风险人工智能系统的提供者承担着最繁重的合规负担，并且需要在系统投放市场或投入使用之前进行符合性评估。

- 可能成为高风险人工智能系统提供者的情形（例如，以自己的名称或商标提供高风险人工智能系统、对高风险人工智能系统进行重大修改或将其用于与原始提供者预期目的不同的用途）。

行动指南

- [ ] 结合附件I和附件III，确定人工智能系统是否落入《人工智能法案》第六条所涵盖的高风险范围。

- [ ] 确定您在价值链中的角色（提供者、部署者、进口商、经销商或第三方供应商）并审查相应的合规义务。

高风险人工智能系统的类别

《人工智能法案》以大量篇幅对高风险人工智能系统作出了规范。这些人工智能系统可能对欧盟个人的健康、安全和基本权利产生显著有害影响。高风险人工智能系统主要分为两类：

a. 附件I所列欧盟统一立法所涵盖的，旨在作为产品或系统的安全部件使用，或者本身就是产品或系统的人工智能系统；并且，这些系统根据相关立法需要进行第三方符合性评估。

b. 其预期用途属于《人工智能法案》附件III所列用例范围的系统。

类别A：附件I所列系统

上述类别（a）所称附件I中的产品安全立法涵盖以下类别：

- 机械
- 玩具
- 娱乐船只和个人水上设备
- 升降机/电梯
- 用于潜在爆炸性环境中的设备和保护系统
- 无线电设备
- 压力设备
- 缆道装置
- 个人防护设备
- 燃烧气体燃料的器具、医疗器械
- 体外诊断医疗器械
- 民航
- 两轮或三轮汽车
- 农林车辆
- 海事设备
- 铁路系统
- 机动车辆及其挂车
- 无人机

请注意，附件I中的法规不仅涵盖属于上述类别的产品，也可能涵盖其他相关产品。例如，《机械法规》涵盖起重配件、可拆卸机械传动装置以及机械。同时，它也是机器人技术的核心法规。机器人技术作为人工智能应用的另一个稳步增长的领域，将与《人工智能法案》及其针对高风险人工智能系统的要求息息相关。

安全部件可实现产品的安全功能，如果其失效或故障会危及人员的健康安全或者财产安全。您应根据附件I中适用的产品安全法规进行评估，以确定人工智能系统是否必须根据该法规接受第三方符合性评估。例如，在《医疗器械法规》中，IIa 类及以上医疗器械必须接受第三方符合性评估程序。如果人工智能系统属于此类医疗器械的安全部件，或者其本身构成此类医疗器械，则根据《人工智能法案》，它属于高风险人工智能系统。

附件I中涵盖的某些立法也使用了“*高风险*”和“*中等风险*”等术语。但是，这些类别与《人工智能法案》规定的高风险分类无关。例如，根据适用的产品安全法规，产品可能被归类为“*中等风险*”，但如果产品必须经过第三方符合性评估，那么作为该产品安全部件的人工智能系统或本身构成此类产品的人工智能系统将根据《人工智能法案》界定为高风险人工智能系统。

类别B：附件III所列系统

构成高风险系统的独立清单目前包含以下类别：

- 生物识别：个人远程生物识别、个人生物识别特征分类或者个人情绪识别。
- 关键基础设施的管理和运行：与关键数字基础设施（例如互联网交换点、DNS 服务、TLD 注册中心、云计算服务、数据中心、内容分发网络、信任服务提供商、电子通信网络或服务）的管理和运行相关，以及在水、气、暖或电力供应过程中，直接保护个人人身完整、健康、安全以及财产安全的安全部件。
- 教育和职业培训：教育和职业培训过程的决策（例如，对学生或申请成为学生的个人进行筛选、评估、考核和监控）。
- 就业和人力资源：招聘和人力资源管理过程中的决策（例如，对员工、其他工人或者申请者进行筛选、评估、考核、晋升、解雇、任务分配和监

控）。

- 基本服务：评估个人（继续）获得公共援助福利的资格（例如医疗保健服务、社会保障津贴、残疾人福利）；评估个人信誉或确定其信用评分（用于检测金融欺诈除外）；在人寿和健康保险方面，对个人进行风险评估和定价；对紧急呼叫进行评估和分类或者作出与调度或确定调度紧急第一反应（Emergency First Response）服务优先次序（例如警察、消防员、医疗救助）有关的决策，以及紧急医疗病人分流。

- 犯罪分析：由执法机关或者代表/支持执法机关的欧盟机构进行的评估，包括：（i）个人成为受害者或（再次）犯罪者的风险；（ii）个人的性格特征和特点；（iii）个人或团体过去的犯罪行为；或（iv）在侦查、调查或起诉刑事犯罪过程中，对个人进行画像分析。

- 证据收集和评估：在调查或起诉刑事犯罪期间，或在申请庇护、签证或居留许可过程中，或涉及相关投诉时，评估证据的可靠性；执法机关、代表/支持执法机关的欧盟机构以及负责移民、庇护或者边境管制的机关使用测谎仪或类似工具。

- 移民识别、移民风险和移民申请评估：在移民、庇护或边境管制的背景下检测、识别或确认个人身份（旅行证件核实除外）；评估打算进入或已经进入欧盟国家领土的个人所带来的风险（例如安全风险、非法移民风险或健康风险），并审查与庇护、签证或居留许可相关的申请及相关投诉。

- 司法行政：协助司法机关或替代性争议解决机构研究和解释事实与法律，并将法律适用于事实。

- 民主进程：影响选举或公民投票的结果，或者影响个人投票行为。

请注意委员会可能会对附件III进行修改（第 7 条）。

> *《人工智能法案》第3(12)条将“预期目的”定义为：“提供者对人工智能系统预期的用途，包括提供者在使用说明、宣传或销售材料、声明以及技术文档中所列明的特定使用场景和条件。”*

例外情形：不足以构成高风险

《人工智能法案》第6(3)条规定，如果人工智能系统的预期用途虽然落入附件III的范围之内（在缺少该条款的情形下会被视为高风险），但是它们不会对自然人的健康、安全或基本权利造成重大损害风险，则它们不应被视为高风险。该条款列出了四项标准，如果满足以下一项或多项标准，则可以主张豁免（第6(3)条以及序言第53条）：

- 人工智能系统旨在执行狭义的程序任务；

  — 示例：将非结构化数据转换为结构化数据的系统，或检测重复文件的人工智能系统

- 人工智能系统旨在改进先前完成的人类活动的结果；

  — 示例：对已起草文档中语言使用的专业度或学术风格进行优化的人工智能系统

- 人工智能系统旨在检测决策模式或先前决策模式的偏差，而不是在未经适当人工审查的情况下取代或影响先前完成的人工评估；

  — 示例：与教师当前的评分模式进行比较，对教师评分中出现的不一致或异常情况进行标记的人工智能系统

- 人工智能系统旨在执行与附件III所列用例相关的评估的准备任务；

  — 示例：用于翻译文档的人工智能系统。

根据《人工智能法案》序言第53条，上述例外情况不适用于任何涉及对自然人进行用户画像的人工智能系统。“用户画像”（profiling）指的是根据欧盟2016/679号条例（《通用数据保护条例》，GDPR）第4(4)条、欧盟2016/680号指令（《数据保护执法指令》，Data Protection Enforcement Directive）第3(4)条或者欧盟2018/1725号条例（《欧盟机构数据保护条例》Data Protection for EU institutions）第3(5)款定义所实施的行为。

决定主张上述例外情形的公司应注意，他们负有证明该人工智能系统是否属于高风险系统的举证责任。根据《人工智能法案》第 6 (3)条进行的评估必须在人工智能系统投放市场或投入服务之前记录下来，并且必须就人工智能系统进行注册（第 49(2)条和第 6(4)条）。此类人工智能系统的提供者必须根据国家主管部门的要求向其提供相关评估文件。

委员会未来会提供指南，具体说明第 6 条的实际履行方式，包括人工智能系统高风险和非高风险用例的综合实例清单（第6(5)条）。委员会还可以通过授权法案增加或修改第 6(3) 条的标准。这些指南预计将在《人工智能法案》生效后六个月内发布。

高风险人工智能系统提供者的义务

《人工智能法案》第三章第二节、第三节和第四节详细列出了高风险人工智能系统提供者和部署者的义务：

高风险人工智能系统提供者的义务	
第2节的要求	确保符合第 2 节的要求（见下文）。
提供者的名称及联系方式	在人工智能系统上（如果无法实现，则在其包装或随附文档上）注明提供者的名称（或其品牌）及其联系方式。
质量管理体系	根据第 17 条的要求建立质量管理体系。（第 17 条要求通过政策、程序和指示等方式将质量管理体系的有关方面记录下来，并列出了需要记录的详细内容）。
文档保存	留存第 18 条中提到的文件，包括： • 技术文档（第 11 条） • 关于质量管理体系的文档（第 17 条） • 经通知机构批准变更的文档（如适用） • 通知机构发布的决定和其他文件 • 欧盟符合性声明（第47条）。
日志	如果人工智能系统在提供者的控制之下，则应留存系统自动生成的日志（第 19 条）。 此类日志的留存期限必须与高风险人工智能系统的预期用途相适应。留存期限至少为六个月（除非任何个人数据保护条款另有规定）。
符合性评估	确保人工智能系统在投放市场或投入使用之前通过第 43 条所规定的符合性评估程序（详见下文）。
符合性声明	起草一份欧盟符合性声明（第 47 条）。 详见下文。
CE 标志	将 CE 标志加贴在高风险人工智能系统上（如果无法实现，则加贴在其包装或随附文档上）。 CE 标志将用以证明高风险人工智能系统符合《人工智能法案》第 48 条的规定。 详见下文。
注册义务	遵守欧盟数据库注册义务（第 49(1) 条）。详见下文。
纠正措施/提供信息	如果人工智能系统不符合《人工智能法案》的合规要求，则应采取必要的纠正措施，或者撤回、禁用或召回该系统。 如果人工智能系统对安全或个人的基本权利构成风险，则应通知市场监督管理机构，并在适用的情况下，通知为该系统颁发合格证书的通知机构（第 79 条）。
合规证明	根据国家主管当局的合理要求，证明人工智能系统符合第 2 节（见上文）规定的各项要求，并提供所有必要的信息和文档。 《人工智能法案》第 21 条更为详尽地规定了与主管当局合作的有关职责。 与国家主管当局共享的任何信息都应被视为机密。
无障碍要求	确保高风险人工智能系统符合以下无障碍要求： • 欧盟2016/2102 号 指令（关于公共部门机构网站和移动应用程序的可访问性）；以及 • 欧盟2019/882号指令（关于产品和服务的可访问性要求）。

高风险人工智能系统提供者的协调标准和符合性评估程序

协调标准

协调标准将在《欧盟官方公报》上公布。如果人工智能系统符合这些标准，则将推定该系统符合《人工智能法案》第三章第二节中对高风险人工智能系统的合规要求。（第 40(1) 条）

协调标准在实践中具有重要意义。根据传统的产品安全法，“制造商”通常会遵循这些标准来证明其符合产品安全法的要求。《人工智能法案》法案也将采取类似做法。

欧盟委员会根据《人工智能法案》第 40(2) 条向标准化机构 CEN/CELENEC发布了一份标准化请求（草案），要求这些机构在2025年4月30日之前起草涵盖第三章第二节要求的协调标准。

符合性评估程序

《人工智能法案》第 43 条规定的高风险人工智能系统符合性评估程序要求提供者证明其符合第三章第二节中关于高风险人工智能系统的合规要求（概述如下）。

附件III 高风险人工智能系统

《人工智能法案》概述了符合性评估的两个主要程序。附件 III 中高风险人工智能系统（即附件 III 第 2 至 8 点所述系统）的提供者必须遵循附件六中规定的内部控制程序，并且无需通知机构的参与。此外，附件 III 第 1 点（生物识别技术）所列的高风险人工智能系统提供者，即便已采用第 40 和 41 条所述的协调标准或共同规范，也必须充分遵循内部控制程序。但是，对于尚未采用上述标准或规范的高风险生物识别系统提供者，则需要在通知机构的参与下进行符合性评估。

附件一 高风险人工智能系统

如果高风险人工智能系统落入附件一第 A 节所列的欧盟统一立法规制的范围内，则应遵循这些法案中规定的符合性评估程序。同时，高风险人工智能系统还应符合《人工智能法案》第三章第2节的合规要求，这些要求会被纳入到评估过程中。此外，附件七的具体规定也应适用。为了确保监管的一致性，这些法案下的通知机构也必须遵守《人工智能法案》的某些要求。

对重大修改进行新的符合性评估

对高风险人工智能系统的重大修改需要进行新的符合性评估。然而，作为系统预定学习过程一部分的变更不被视为重大修改。

重点关注第8-15条; 对高风险人工智能系统的要求	
合规要求(第8条)	第8条明确指出，高风险人工智能系统在其全生命周期内必须满足技术和组织要求（第9-15条），并结合其预期用途和技术现状予以落实。在这一过程中，应优先考虑对人类产生重大影响的要求。若无法找到合适的权衡方案，则不应部署该人工智能系统。
风险管理(第9条)	第9条要求人工智能系统的提供者建立风险管理体系。该体系是一个持续性流程，旨在识别、分析并降低可预见的风险。具体措施包括设计风险缓解措施、实施管控机制以及为用户提供信息和培训。所有采取的措施都必须进行记录，并且高风险人工智能系统应在适当时机接受测试，以确保其性能的稳定一致性。
数据治理(第10条)	健全的数据治理是高风险人工智能系统技术和组织要求的核心要素。为保障系统的正常运行与安全性，必须确保训练、验证和测试数据集的质量——数据集应具有高质量、代表性，并尽可能做到无误且完整。提供者还需采取措施，消除数据集中可能导致禁止性歧视的偏见，包括在特定条件下处理特殊类别的个人数据。此外，可借助经认证的第三方服务进行数据完整性验证，以证明符合《人工智能法案》的数据治理要求。
技术文档和记录保存(第11条和第12条)	第11条和第12条要求在整个系统生命周期内提供详细的技术文档和记录保存日志。提供者必须在系统部署前准备好这些文档，并定期更新。技术文档应涵盖系统的各个方面，包括其特性、算法、数据、训练、测试、验证和风险管理。高风险人工智能系统还应自动记录使用日志，以提供可追溯性并识别潜在风险或需要的修改。

透明度与信息提供(第13条)	第13条要求为高风险人工智能系统的部署者提供清晰、全面的说明。这些说明应使部署者能够正确理解和使用系统的输出。系统的决策过程必须是可理解的，并且需要提供有关其身份、特性、限制、目的、准确性、风险、能力、监督、维护和预期使用寿命的详细信息。所有文档都应根据目标部署者的需要和知识水平进行定制。
人工监督 (第14条)	人工监督措施必须防止或最小化对健康、安全和权利的风险。这些措施必须与系统的风险和自主性水平相称。必要时，人工操作员应能够干预系统的决策。 监督可以通过以下方式实现： • 系统内置的约束机制以及对人工操作员的响应能力。 • 提供者为部署者制定的措施，以帮助他们做出知情的、自主的决策。 • 监督方式可以根据应用的风险程度，选择HITL（Human-in-the-loop）[1]、HOTL（Human-on-the-loop）[2]或HIC（Human-in-command）[3]等模式。
准确性、鲁棒性与网络安全（第15条）	第15条规定，高风险人工智能系统必须达到适当的准确性、鲁棒性和网络安全水平。准确性指标包括最小化预测误差；鲁棒性指标是确保系统能够处理错误和不稳定情形；网络安全措施则旨在防范未经授权的系统篡改。对于受《欧盟网络弹性法案》（Cyber Resilience Act）约束的相关人工智能系统，其合规性可通过该法案进行验证。

高风险人工智能系统部署者的义务

《人工智能法案》第26条对高风险人工智能系统部署者的义务作出规定：

技术和组织措施

部署者应采取适当的技术和组织措施，确保按照系统所附的使用说明使用高风险人工智能系统。

人工监督

部署者需指派有专业能力、接受过相应培训且拥有相应权限的自然人进行人工监督，并为其提供必要的支持。

输入数据质量

若部署者对输入数据具有控制权，则应确保输入数据与高风险人工智能系统的预期用途相关且具有充分的代表性。换言之，这一原则明确了部署者对输入数据质量的责任。

高风险人工智能系统的监测

部署者应根据系统所附的使用说明，对高风险人工智能系统的运行情况进行监测。

部署者需按照第72条关于售后活动的规定，向提供者通知相关信息。若部署者根据第79条第1款识别出相关风险，应立即通知提供者，并随后通知进口商或分销商以及相关市场监督机构，并暂停使用该系统。若发现严重事件，部署者也应立即通知提供者、进口商或分销商以及相关市场监督机构。

日志记录

高风险人工智能系统的部署者应当保存由该系统自动生成的日志记录，前提是这些日志处于其控制范围内，且保存期限应与系统的预期用途相匹配。该期限不得少于六个月，除非适用的欧盟或国家法律另有规定，尤其是在个人数据保护方面。

1. 人类积极参与并直接介入系统的决策过程。
2. 人类处于监督角色，虽然不直接参与决策，但随时准备介入干预，以确保系统运行符合预期。
3. 人类处于主导地位，系统完全由人类控制并下达指令。

向职工代表提供信息

作为雇主的高风险人工智能系统部署者，必须向职工代表及受影响的职工告知，他们将受到高风险人工智能系统使用的影响。

公共机关部署者

作为公共机关或欧盟机构、机关、办事处或部门的高风险人工智能系统部署者，应当遵守《人工智能法案》第49条规定的欧盟数据库注册义务。

数据保护影响评估（DPIA）

如果高风险人工智能系统的提供者依据欧盟2016/679号条例（《通用数据保护条例》，GDPR）第35条或欧盟2016/680号指令（《数据保护执法指令》第27条需要开展数据保护影响评估，则必须利用《人工智能法案》第13条提供的信息。

刑事犯罪调查——高风险人工智能系统用于事后远程生物特征识别

在不影响欧盟2016/680号指令（《数据保护执法指令》的前提下，针对涉嫌或已被判定刑事犯罪的人员进行目标搜查时，若需部署高风险人工智能系统用于事后远程生物识别，相关主体必须事先或在不晚于48小时的合理时间内，向司法机关或行政机关申请使用授权。

高风险人工智能系统的基本权利影响评估（FRIA）

在部署《人工智能法案》第6条第2款所述的高风险人工智能系统（即《人工智能法案》附件III中详细列出的高风险人工智能系统）之前，以下主体必须进行基本权利影响评估（FRIA）：

I. 公共机构；

II. 由公法设立的机构；

III. 提供公共服务的私人实体，且这些主体均部署高风险人工智能系统用于以下用途：

a. 评估自然人的信用状况或确定其信用评分（不包括用于检测金融欺诈的人工智能系统）；

b. 在人寿和健康保险中对自然人进行风险评估和定价。涉及关键基础设施的高风险人工智能系统不在此列。

评估内容包括以下方面：

- 描述高风险人工智能系统将如何按照其预期目的被使用；
- 明确该系统计划使用的时间范围和频率；
- 识别可能在特定情境下受到系统使用影响的自然人或群体类别；
- 结合提供者根据《人工智能法案》第13条提供的信息，分析可能对上述自然人或群体类别造成伤害影响的具体风险；
- 描述根据使用说明，人类监督措施的实行情况；
- 包括内部治理安排和投诉机制在内的，针对上述风险的应对措施

高风险人工智能系统其他相关方的义务

《人工智能法案》中关于高风险系统的大部分义务主要针对提供者和部署者。然而，也有部分义务适用于其他相关方，包括高风险人工智能系统的进口商、分销商，以及为高风险人工智能系统提供系统、工具、服务、组件或流程的供应商（这些系统、工具、服务、组件或流程被用于或整合到高风险人工智能系统中）。供应商提供的服务可能包括模型（重新）训练、测试与评估，以及与软件的集成（序言第88条）。根据第25条第4款，这些义务不适用于以自由开源许可提供相关产品或服务的供应商。此外，根据《人工智能法案》，除初始人工智能系统的提供者外，其他方也可能被指定为高风险人工智能系统的提供者。

进口商、分销商和供应商的义务

第23条、第24条和第25条明确了进口商、分销商和供应商的义务：

进口商 （第23条）	分销商 （第24条）	供应商 （第25条）
验证义务：将系统投放市场之前，验证提供者是否切实履行以下义务: • 已实施符合性评估程序; • 已编制技术文件; • 已加贴CE标志并附上欧盟符合性声明； • 已指定授权代表。	验证义务：在使系统在市场上可用之前，需验证以下内容: • 是否加贴CE标志； • 是否附有欧盟符合性声明副本及使用说明； • 提供者和进口商（如适用）是否已履行其各自义务。	提供协助义务: 通过书面协议，明确必要的信息、能力、技术访问权限以及其他基于公认的最新技术水平的协助内容，以确保高风险人工智能系统的提供者能够充分履行其义务。 此外，人工智能办公室/委员会可制定并推荐高风险人工智能系统提供者与其第三方供应商之间的自愿性示范合同条款（参见第25条第4款和序言第90条）。
风险警示：当系统存在对人员健康、安全或基本权利构成风险的情形时[4]，应立即通知提供者、授权代表以及市场监管机构。	风险警示：如果分销商认为或有合理理由怀疑系统不符合第2节所规定的要求，则不得将该系统投放市场，直至其符合相关要求。若系统存在对人员健康、安全或基本权利构成风险的情形，分销商应立即通知系统提供者或进口商以及相关主管机构，并详细说明不符合的具体情况以及已采取的纠正措施。	
注意：确保储存或运输条件不会危及第 2 部分的要求。	注意：确保储存或运输条件不会危及第 2 部分的要求。	
与主管当局合作：应主管当局的合理要求，向其提供所有必要的信息/文件，包括技术文件，以证明系统符合要求，并与主管当局合作采取任何与系统有关的行动。	与主管当局合作：应主管当局的合理要求，向其提供所有必要的信息/文件，包括技术文件，以证明系统符合要求，并与主管当局合作采取任何与系统有关的行动。	
记录保存：在系统投放市场/投入使用后十年内，保留以下文件的副本：通知机构颁发的证书（如果是第三方符合性评估）、使用说明和欧盟合格声明。 联系方式：在系统及其包装或随附文件上注明进口商的名称、注册商号或注册商标以及联系地址。	纠正措施：当分销商认为或有理由认为系统不符合第 2 节规定的要求时，采取必要的纠正措施，使系统符合要求，或撤回或召回该系统，或确保提供商、进口商或任何相关运营商根据情况采取这些纠正措施。	

4. 此处的“风险”是指：“可能对人员的健康和安全产生不利影响，且这种影响超出了与其预期用途或在正常使用或合理可预见的使用条件下（包括使用期限以及适用的投入使用、安装和维护要求）被认为合理和可接受的程度”（《人工智能法案》第79条第1款，结合《欧盟市场监督条例》（EU）2019/1020第3条第19款）。

成为他人（高风险）人工智能系统的提供商

第 25(1) 条规定，在下列情况下，某人将被视为高风险人工智能系统的提供者，即使此人最初并非该人工智能系统的提供者：

- 在已经投放市场或投入使用的高风险人工智能系统上使用自己的名称或商标；
- 对现有的高风险人工智能系统进行重大修改[5]，使其仍然具有高风险；和/或
- 修改目前不属于高风险的人工智能系统的预期用途，使其成为高风险系统。

如果发生上述三种情况中的任何一种，原提供者将不再被视为（新的或新使用的）人工智能系统的提供者。实践中经常发生的一种导致提供者角色转换的可能是，部署者以属于第 6 条（以及附件 I 和 III ）规定的高风险类别的方式部署通用人工智能系统。因此，如果某人以高风险的方式部署通用人工智能系统，则该部署者将承担提供者的责任。

新提供者将承担高风险人工智能系统提供商的所有义务。原提供者有义务与新提供者密切合作，提供必要的信息，并向新提供者提供合理预期的技术访问权限和其他协助，使该系统符合《人工智能法案》（第 25(2) 条）。但是，如果原提供者“明确规定”不得将人工智能系统更改为高风险人工智能系统（第 25(2) 条）或“明确排除将人工智能系统更改为高风险人工智能系统”（第 86 条），例如在适用合同中禁止为高风险目的部署人工智能系统，则原提供商没有义务这样做。如果不禁止高风险部署人工智能系统，则合作义务适用，但不影响遵守和保护知识产权、机密商业信息和商业秘密的需要（第 25(5) 条）。因此，原始提供者不必在损害其自身知识产权或商业秘密的情况下提供帮助（第 88 条）。

委员会将就本第 25 条所述要求和义务的适用提供指导方针 (第 96(1)(a) 条)。

我在哪里可以找到这个？

高风险系统的范围	第6条，附件 I	序言第46-63条，附件 III
对高风险人工智能系统提供者的要求	第8-22、43、47-49条	序言第64-83、123-128、147、131条
高风险人工智能系统部署者的要求	第26、27条	序言第91-96条
高风险人工智能系统进口商的要求	第23条	序言第83条
高风险人工智能系统分销商的要求	第24条	序言第83条
高风险系统第三方供应商的要求	第25条	序言第83-90条
标准	第40、41条	序言第121条
合格评定程序	第28条	序言第149条

5. 第 3(23) 条将“重大修改”定义为“人工智能系统在投放市场或投入使用后发生的变更，这种变更在提供商进行的初始合格评定中未预见或未计划，并且因此影响了人工智能系统对第 III 章第 2 节规定的要求的符合性，或导致对人工智能系统评估的预期用途进行了修改”。委员会将就实质性修改相关规定的实际实施提供进一步的指导（第 96(1)(c) 条）。第 84 条还规定，基于新立法框架的某些联盟协调立法（例如《医疗器械条例》）中制定的条款应继续适用。例如，《医疗器械条例》第 16(2) 条规定，某些更改不应是可能影响设备符合适用要求的修改，这些规定应继续适用于高风险人工智能系统，即《医疗器械条例》所定义的医疗器械。

通用人工智能模型

概览

- 通用人工智能模型是多功能的人工智能组件，具备极高的任务通用性，特别是当前的生成式人工智能模型。
- 对通用人工智能模型进行微调和修改可能会产生新的通用人工智能模型。
- 通用人工智能模型提供者需向人工智能办公室和主管当局履行多项透明度义务，同时也需要对意图将其模型集成到人工智能系统中的系统提供者履行义务。
- 具有系统性风险的通用人工智能模型，即目前用途最广泛、功能最强大的模型，需满足更高要求，包括透明度、安全性、风险评估和事件管理等义务。如何对具有系统性风险的通用人工智能模型进行分类应成为通用人工智能模型提供者关注的重点领域。
- 行为准则的制定和发布将帮助通用人工智能模型提供者明确具体的技术和组织措施，以履行其义务。
- 通用人工智能模型相关条款将于 2025 年 8 月 2 日起生效。

行动指南

- 熟悉概念：了解通用人工智能模型、通用人工智能系统、人工智能系统和高风险人工智能系统的概念及其相互关系。这是评估贵公司使用或销售的系统以及做出法律判断的基础。
- 通用人工智能模型提供者：应进行全面的治理审查并做出必要调整以确保合规——通用人工智能模型提供者的义务是《人工智能法案》中最严格的部分之一。
- 通用人工智能模型提供者：需进行全面的知识产权法律评估——通用人工智能模型的监管与知识产权法律紧密相关，特别是在版权政策及训练数据义务方面。
- 通用人工智能模型提供者：应持续密切关注“系统性风险”的阈值，因为这些阈值可能会随着时间的推移通过授权法案进行调整。
- 通用人工智能模型提供者：应密切关注行为准则的制定和发布，其中将包括具体的技术细节，指导通用人工智能模型提供者如何在实践中履行义务。订阅 Bird & Bird 的 Connected newsletter，了解最新动态！

通用人工智能模型的背景及相关性

在《人工智能法案》立法过程中，关于通用人工智能的监管一直是一个重要的议题。《人工智能法案》的初稿（欧盟委员会 2021 年 4 月的提案）基于的理解为：每个人工智能系统都是有特定目的，并且此目的与特定的风险潜能相关。这种分类未考虑到基础模型，这些模型通过广泛的数据训练，能供应用于各种场景。这些人工智能模型不符合《人工智能法案》初稿中基于风险的分类方案。因此，分类需要扩展，通过加入新的类别，以考虑此类基础模型的特定能力和风险。2023 年夏季，“基础模型（foundation model）”（后更名为通用人工智能（general-purpose AI））被加入到当时的《人工智能法案》草案中。

《人工智能法案》中关于通用人工智能模型监管的章节具有重要意义，主要有以下两个原因：

- 首先，它涉及生成式人工智能，这是目前在商业环境开辟最具吸引力的新机遇的人工智能子类别，并涵盖了大多数企业用例；
- 其次，《人工智能法案》对通用人工智能的要求与对高风险人工智能系统的要求一样，是《人工智能法案》中最严格的，需要企业在实施时极其审慎。

这一重要意义仅因所有要求仅针对提供者而非部署者而略有减弱。

术语与通用人工智能价值链

通用人工智能模型和通用人工智能系统

《人工智能法案》第3（63）条概述了通用人工智能模型的特点，强调了其在各种任务中的多功能性和能力。

序言第 98 条强调了两个关键指标：

1. 至少具有十亿个参数；
2. 使用大量数据进行自我监督训练。

这些模型的特点是能够集成到各种下游系统或应用中并在其中发挥作用。通常，通用人工智能模型会使用自我监督等方法进行大规模的数据训练。序言第 99 条进一步指出，大型生成式人工智能模型，例如大语言模型（LLM）或扩散模型（Diffusion Models），是通用人工智能模型的典型示例。

序言第 97 条明确指出，虽然通用人工智能模型是人工智能系统的重要组件，但它们本身并不是人工智能系统。需要用户界面等其他组件才能将通用人工智能模型转化为完全可操作的人工智能系统。通用人工智能系统是基于通用人工智能模型构建的人工智能系统，在各种任务中保持其多功能性（第 3(66) 条和序言第 100 条）。举例而言，仅执行翻译任务的系统可能并不被视为通用人工智能系统。

通用人工智能系统与高风险人工智能系统

序言第 85 条强调，通用人工智能系统由于其多功能性，可能成为高风险人工智能系统或高风险人工智能系统的组成部分。通用人工智能系统的提供者必须与高风险人工智能系统的提供者密切合作，以确保遵守《人工智能法案》的要求并在人工智能价值链中公平分配责任（有关高风险系统的更多信息，请参阅第 4 章）。

通用人工智能模型的修改与微调

对通用人工智能模型进行修改或微调，即将新的专业化训练数据输入模型以提升特定任务的表现，并不会将该模型转变为通用人工智能系统。它仍然是一个缺乏用户界面的抽象模型。相反，此类操作创建的是一个修改或微调后的通用人工智能模型。序言第 97 条和第 109 条规定，修改或微调通用人工智能模型的提供者仅对其所做的更改承担有限的义务，包括提供技术文档或所用训练数据的摘要。

通用人工智能模型提供者的义务

通用人工智能模型提供者在将模型投放市场、或将其与自有人工智能系统集成后投放市场或投入服务时，应当履行以下义务：

a. 编制并不断更新技术文档，包括模型描述及开发过程（如培训、测试和验证）的信息，以便应要求向人工智能办公室和主管当局提供（第 53(1)(a) 条）——具体最低限度信息要求详见附件 XI；

b. 编制、不断更新并向下游人工智能系统提供者（即希望将其人工智能系统与通用人工智能模型相结合的主体）提供必要的信息和文档，确保其了解该模型的特点并履行自身义务（第 53(1)(b) 条）——具体最低限度信息要求详见附件XII；提供者可在信息共享与保护商业机密信息和商业秘密的需要之间取得平衡；

c. 制定政策以遵守欧盟版权及相关权利的法规（第 53(1)(c) 条），尊重欧盟第 2019/790 号指令（《关于数字单一市场的版权和相关权利指令》）中第 4(3) 条规定的文本和数据挖掘的选择退

出权 。《人工智能法案》未具体规定政策中需涵盖的其他事项；

d. 起草并公开模型训练数据的全面摘要（第 53(1)(d) 条）——人工智能办公室的任务将为此提供模板；正如序言第 107 条所言，摘要应列出主要数据集、数据库或数据档案等，以便相关方行使其权利；

e. 在有关当局行使《人工智能法案》赋予的权力时，予以配合（第 53(3) 条）；

f. 如果提供者设立于欧盟以外：则需在欧盟内指定一名授权代表（第 54(1) 条）。

如果提供者根据免费和开源许可发布通用人工智能模型并公开相关信息，则不必履行上述 (a)、(b) 和 (f) 项的要求——除非该通用人工智能模型被认定为存在系统性风险（第 53(2) 条和第 54(6) 条）。

具有系统性风险的通用人工智能模型

资格标准

《人工智能法案》为具有“系统性风险”的通用人工智能模型（如合理可预见地将带来重大事故，造成关键行业的中断，导致关于公共健康和安全、公共和经济安全、民主进程地严重后果，或虚假或歧视性内容扩散等负面影响的模型）引入了特别的增强义务（序言第 110 条）。

根据《人工智能法案》第51(1)条，如果通用人工智能模型符合以下两个条件之一，则该模型将被归类为具有系统性风险的通用人工智能模型：(a) 具有基于技术工具和方法评估的“高影响力（high impact capabilities）”，或 (b) 经欧盟委员会指定，具有与 (a) 项相当的能力或影响力，具体标准参见《人工智能法案》附件 XIII。这些标准主要包括模型的参数数量、数据集的质量或规模、用于训练的计算量、模型对欧洲市场的影响、欧盟注册用户数量等。

此外，如果模型训练过程中使用的浮点运算超过 10^25 次，即具备强大计算能力，则该模型将被视为具有“高影响力”（第 51(2) 条）。在本指南发布时，只有少数大语言模型似乎符合这一门槛。

《人工智能法案》第 52 条规定了分类程序。值得注意的是，符合系统性风险分类条件的通用人工智能模型的提供者在不迟于满足此条件或得知将满足此条件后的两周内立即通知欧盟委员会。提供者可以提出论据，证明尽管满足要求，但其模型并不构成系统性风险。如果此类理由被欧盟委员会驳回，则该模型将被视为存在系统性风险。若提供者存在“合理请求（reasoned request）”，欧盟委员会可决定重新评估分类（第 52(5) 条）。

欧盟委员会将发布并定期更新具有系统性风险的通用人工智能模型清单（第 52(6) 条）。

具有系统性风险的通用人工智能模型提供者的义务

除了适用于所有通用人工智能模型提供者的一般要求外，《人工智能法案》还对具有系统性风险的通用人工智能模型提供者施加了额外的增强义务（第 53(1) 条和第 55(1) 条）。这些义务适用于模型投放市场之前及其整个生命周期内，涉及：

- 模型评估；
- 系统性风险评估与缓解；
- 事件管理和报告；
- 提升网络安全保护水平；以及
- 扩展的技术文档。

透明度义务

概览

《人工智能法案》根据风险等级对人工智能系统进行了分类，对高风险类的透明度要求更高。高风险人工智能系统在投放市场或投入使用之前必须确保透明度。有关高风险人工智能系统透明度要求的更多详细信息，请参阅本指南第 4 章。

此外，《人工智能法案》第50条规定了特定类型产品的透明度要求，要求提供者或部署者向个人提供足够的信息。

- 免责声明：旨在直接与个人互动的人工智能系统的提供者需要通过设计和开发，让个人意识到其正在与人工智能系统互动。
- 标记要求：人工智能系统提供者必须对人工智能生成的内容（音频、图像、视频、文本）进行标记，将此类内容同人类生成的内容区分开来。
- 深度伪造标记：对人工智能生成的，与真实实体相似且可能误导人们信以为真的人工智能生成内容（图像、音频、视频）必须予以标记。
- 情感识别系统/生物特征分类系统：人工智能系统的部署者应使个人了解这些系统的运行。

《人工智能法案》的透明度义务与欧盟的其他监管框架保持一致。具体而言，《通用数据保护条例》（GDPR）和《人工智能法案》在透明度要求方面有部分重叠，不过后者更偏技术性。

行动指南

提供者

- [ ] 添加标记：确保用机器可读的格式标记人工智能生成内容。
- [ ] 添加免责声明：确保在用于与个人直接互动的人工智能系统中添加适当的免责声明。

部署者

- [ ] 深度伪造：以清晰可辨的方式标注“Deepfake”表明所涉内容系人工创作或操纵。
- [ ] 情感识别系统/生物特征分类系统：让个人意识到该系统正在运行。

一般透明度义务

《人工智能法案》承认透明度在人工智能系统使用中的重要性。应使个人了解人工智能系统的设计和使用，公司和公权力机关应对其决策负责。透明度对于建立公众对人工智能系统的信任以及确保其负责任的部署也至关重要。

透明度还强化了“人工智能素养”这一更广泛的概念，提高人们对人工智能带来的机遇和风险以及可能造成的危害的认知，特别是以下群体：

- 相关个人能够更好地了解其在人工智能背景下的权利，以及
- 部署者能够以知情的方式部署人工智能系统。

提供者以及某些情况下的部署者有各自的透明度要求。《人工智能法案》根据风险等级对人工智能系统进行分类，风险等级越高，透明度要求就越高。

特定类型产品的透明度要求如下所述。

提供者义务：

聊天机器人（《人工智能法案》第50(1)条）

《人工智能法案》第50(1)条规定，人工智能系统提供者需确保用于与个人直接互动的人工系统的设计和开发能够让相关个人意识到其正在与人工智能系统互动。

- 目标受众：在履行透明度义务时，提供者不仅应确定预期的目标受众，还应确定可能会看到免责声明的更广泛的潜在目标受众。如果人工智能系统还会用于与因年龄或失能而属于弱势群体的受众互动，也要考虑这类群体的个人特征。预期或潜在目标受众对可访问性的考虑因素有重大影响。
- 形式：在实践中，提供者可以设计不同形式（例如头像、图标或界面）的免责声明，只要能明确表明个人正在与人工智能系统交互。

豁免

- 显而易见：考虑到人工智能系统使用的情形和场景，如果对于一个合理知情、善于观察和谨慎的个人来说，其正在与人工智能系统互动是显而易见的，则该人工智能系统可免除该透明度要求。
- 合法用途：经法律允许用于检测、预防、调查或起诉犯罪活动的人工智能系统，在对第三方权利和自由采取了适当保障措施的情况下，可免除上述透明度要求，除非这些系统可供公众举报犯罪行为。

人工智能生成内容的标记（《人工智能法案》第50(2)条）

《人工智能法案》第50(2)条规定，人工智能系统（包括通用人工智能系统）的提供商必须适当标记音频、图像、视频或文本等合成内容。引言133项解释了其基本原理：随着人工智能技术的进步，人工智能生成的合成内容与人类生成的内容越来越难以区分，带来了错误信息、操纵、欺诈、冒充和欺骗消费者的风险。

标记义务

- 标记：只有人工智能系统的提供者被要求标记人工智能生成内容。此要求不适用于内容的部署者或其他用户。
- 格式：必须以机器可读的格式对输出进行标记，以表明是人工生成或操纵的。
- 技术标准：标记应有效、可互操作、稳定且可靠。提供者需要考虑内容类型、实施成本和当前的技术标准。

标记方法：

- 水印：添加可见水印很简单——但是很容易被基础的编辑工具删除，而隐形水印需要用专门的软件才能检测和删除。
- 元数据：提供有关文件创建和来源的信息，但是可以使用文件编辑工具轻松更改或删除。
- 算法指纹：人工智能模型在其生成的内容中留下独特的痕迹或异常现象。例如，人工智能生成的图像在纹理或图案上可能会有轻微的扭曲，人工智能创建的音频文件可能会显示不自然的停顿或音调变化。
- 加密签名：使用加密方法嵌入的数字签名，例如用于验证内容真实性的加密哈希值。即使数据的微小变化也会导致不同的哈希值，从而确保便捷地验证内容是否被修改。

目前，有许多工具和措施可用于管理和检测人工智能生成的内容。有些平台使用深度伪造检测软件来分析算法模式和嵌入的元数据，而有些平台则依靠元数据和加密哈希值来验证内容来源。例如，平台可能使用语音分析工具来检测合成音频，或者使用区块链技术来追踪数字艺术品的来源和修改。

豁免

- 编辑辅助：主要为日常编辑任务提供支持或不会对原始输入数据进行重大改变的人工智能系统无需承担标记义务。

- 合法用途：经授权用于检测、预防、调查或起诉犯罪活动的人工智能系统也不受标记要求的约束。

部署者义务：

情感识别/生物特征分类系统（《人工智能法案》第50(3)条）

《人工智能法案》第50(3)条规定了对部署者的具体透明度要求：

- 情感识别系统：用于根据自然人的生物特征数据（例如面部表情等非语言信号）识别或推断自然人的情感或意图的人工智能系统。

或者

- 生物特征分类系统：用于根据自然人的生物特征数据将其归类为特定类别的人工智能系统。所述特定类别可能与性别、年龄、头发颜色、眼睛颜色、纹身、个人特征、民族血统、个人喜好和兴趣等方面相关。

关于在何种情况下禁止使用情感识别系统或生物特征分类系统的更多详细信息，请参阅本指南第4章。

当这些系统被允许使用时，部署者必须向接触这些系统的自然人告知有关系统的使用情况。特别是，当个人接触人工智能系统时应向其告知，人工智能系统通过处理他们的生物特征数据，可以识别或推断他们的情感或意图，或者将其归入特定类别。

豁免

- 合法用途：获准用于检测、预防或调查犯罪活动的人工智能系统，在对第三方权利和自由采取了适当保障措施的情况下，并且符合联盟法律，可免于上述要求。

- 辅助用途的生物特征分类系统：用于辅助其他商业服务且由于客观技术原因而必不可少的人工智能系统可免于上述要求。

关于应当提供的信息的范围，目前尚无明确的指导方针。部署者在使用这些系统时，除了有关处理的法律依据的要求外，还要根据GDPR和(EU) 2018/1725号指令以及(EU) 2016/680号指令（如适用）处理个人数据。这意味着这些规则也构成部署者作为控制者的单独的透明度义务。在这种情况下，仍然应当根据GDPR第13条和第14条的要求向个人告知其数据处理情况。对于任何自动化处理，控制者还应解释其决策背后的逻辑。就人工智能系统而言，这可以包含在可解释性声明中。可解释性声明提供非技术性的解释，说明该组织为何使用人工智能、人工智能是如何开发的以及人工智能是如何运作和使用的。

深度伪造（《人工智能法案》第50(4)条）

《人工智能法案》第50(4)条规定了被称为 *"Deepfakes"* 的内容的具体标签要求。当使用人工智能系统被用来生成或操纵内容时，这些义务对于保证透明度至关重要。

深度伪造的定义（《人工智能法案》第3(60)条）

部署者使用人工智能创建以下内容：

- 生成或操纵图像、音频或视频；

- 与真实的人、物体、地点、实体或事件十分相似；和

- 可能会误导人们相信内容是可靠的或真实的。

深度伪造示例：

- 模仿公司高管的深度伪造视频通话，诱骗员工转移大笔资金。

- 人工智能生成的政治人物音频，利用机器人电话发布误导选民的选举日期信息。

- 深度伪造视频广告冒充政治人物，操纵社交媒体上的公众舆论。

- 使用深度伪造技术冒充知名人士进行虚假Zoom采访。

- 数字化身传递虚假的新闻报道来欺骗观众。

标签要求

《人工智能法案》规定，任何由人工智能系统生成或操纵的内容都必须清晰可辨地加上标签，以表明其是人工创造或操纵的。这一要求旨在确保透明度，防止公众被此类内容误导。关于如何标记内容，目前尚无明确的指导方针。这个问题可能会在将来的行为准则中得到解决。

应当采用水印、元数据标识、指纹或其他技术来表明内容的人工性质（见序言133项）。至关重要的是，这些标签必须易于、即时且持续地被观众看到。例如，就视频而言，可以使用前置标签或持久水印来有效满足这些要求。

豁免

根据《人工智能法案》第50(4)条的规定，标签要求有若干豁免和例外：

- 对于艺术、创意、讽刺、虚构或类似作品，透明度要求更为宽松。这些作品包括例如人工智能生成的电影或恶搞作品、数字艺术展览和人工智能生成的音乐视频。在这些情况下，有义务以不破坏观众体验的方式表明人工智能的参与，可以通过不易察觉的水印、简短的音频免责声明或数字平台上的描述文字注释来实现。

- 如果人工智能系统的使用基于法律的授权，用于检测、预防、调查或起诉刑事犯罪的目的，则不适用给人工智能生成内容加标签的义务。

- 如果人工智能生成的内容已经经过人工审核或编辑控制，并且由自然人或法人对出版物承担编辑责任，则可能不适用标签义务。这意味着，如果人工审核并批准了人工智能生成的内容，确保其准确性和完整性，则标签要求可以不那么严格。这一例外承认了人工监督在维护人工智能生成内容的质量和可靠性方面的作用。

高风险人工智能系统的透明度义务

第50(6)条解释称，此处概述的透明度义务与其他监管要求同时适用，既不取代也不削减第三章规定的义务或者欧盟或国家立法规定的其他透明度要求。

请参阅本指南第4章了解更多详细信息。

时间和形式

履行第50条规定之透明度义务所需的所有信息必须提供给有关个人：

- 以清晰可辨的方式；

- 不迟于所述个人第一次与人工智能系统互动或接触时；并且

- 符合适用的可及性要求。

可及性要求意味着信息应当提供给不同受众（包括残障人士）无障碍获取。在实践中，这可能意味着，根据具体情况，免责声明或其他标记方法不仅必须以书面形式显示，还必须以听觉和（音频）视觉形式显示。

另一个需要考虑的方面是，应当向个人提供清晰、充分但不过量的信息。

国家层面的透明度义务和行为准则

根据《人工智能法案》第50(6)条，《人工智能法案》第50(1)-(4)条概述的透明度义务旨在与其他监管要求共存，既不取代也不削减第三章或者欧盟或国家法律规定的其他透明度要求。

根据《人工智能法案》第50(7)条，人工智能办公室负责促进和推动行为准则的制定，以支持在欧盟层面有效履行《人工智能法案》第50(1)-(4)条规定的透明度义务。这些准则旨在阐明检测和标记人工智能生成内容的方法，加强整个价值链中的合作，确保公众能够清楚地区分人类创建的内容和人工智能生成的内容（引言135项）。

与其他监管框架的关系

- 《人工智能法案》第50(2)-(4)条规定的标记义务支持《数字服务法案》（DSA）对超大型线上平台（VLOP）和搜索引擎（VLOS）的要求，以识别和降低与传播深度伪造有关的风险（DSA第33条及以下条款）。如果人工智能提供者与VLOP或VLOS分开，平台就能通过这些标记更有效地识别人工智能生成的内容。相反，如果VLOP或VLOS也是人工智能提供者，其根据DSA承担的义务在《人工智能法案》进一步详细阐述和强化。

- 深度伪造的透明度法规将与欧洲有关误导性广告的指导方针（参见《不公平商业行为指令》）以及有关深度伪造的国家刑事条款相一致。

- 《人工智能法案》的透明度义务也支持并且补充了(EU) 2016/679号指令中的透明度要求。但是，如果在人工智能生命周期的所有不同阶段（例如在开发、测试或部署人工智能技术）使用人工智能技术时处理个人数据，则GDPR的透明度要求适用，并且适用于控制者。人工智能工具的开发者和提供者并不总是扮演这样的角色。在这种情况下，他们可能仍有义务向控制者提供具体信息，以便后者能够履行其义务。

人工智能监管沙盒

概览

- 《人工智能法案》支持建立“*人工智能监管沙盒*”，以提供一个受控环境，在一定期限内对即将投放市场的创新人工智能系统进行测试。
- 这项制度旨在鼓励人工智能提供者（或潜在提供者）在监管机构的监督下试测创新产品。法案中有具体的激励措施鼓励中小企业和初创企业参与。
- 每个成员国必须在2026年8月2日之前建立至少一个人工智能监管沙盒，也可以与其他成员国合作完成。
- 预计欧盟委员会将会采取行动，对人工智能监管沙盒的建立、运行和监督做出详细安排。
- 《人工智能法案》还规定在监管沙盒内部和外部对人工智能系统进行“*真实世界*”测试，前提是遵守特定保护参与者的条件。
- 人工智能监管沙盒和真实世界测试相关制度应在整个欧盟范围内保持一致。不过，各国可能采取不同的做法，导致提供者有可能“*挑选法域*”。

行动指南

- 参与人工智能监管沙盒和真实世界测试是自愿的。如果人工智能提供者打算参与沙盒或真实世界测试，则需要熟悉《人工智能法案》的相关规定，并且应当关注有关这些主题的进一步公告和指引，包括欧盟委员会将在适当时候指定的人工智能监管沙盒的详细安排。
- 有关主体应考虑要在哪些国家/地区对人工智能服务/产品进行测试。虽然《人工智能法案》旨在建立一个统一的制度，但各国之间仍可能存在差异，导致某些成员国成为您的优选地。
- 一旦决定参与人工智能监管沙盒，则有关主体需要制定沙盒计划，遵循相关国家主管部门的指引和监督。如果决定进行真实世界测试，还需要制定测试计划并提交有关市场监管机构批准。
- 成功完成人工智能监管沙盒流程后，有关主体应当从相关国家主管部门取得退出报告书，这可能有助于加快相关人工智能产品/服务的合格性评估程序。

人工智能监管沙盒

《人工智能法案》允许创建“监管沙盒”，以提供一个受控环境，以便在一定期限内对即将投放市场或以其他方式投入使用的创新人工智能系统进行测试。人工智能监管沙盒制度的目标包括：

- 促进人工智能创新，同时确保创新人工智能系统符合《人工智能法案》；
- 为创新者提供更高的法律确定性；
- 加深国家主管部门对使用人工智能的机会、风险和影响的了解；
- 支持合作和分享最佳实践；以及
- 加快市场准入，包括消除中小企业和初创企业所面临的障碍。

《人工智能法案》规定的监管沙盒是什么？

《人工智能法案》将“*人工智能监管沙盒*”定义为：

“一个由主管部门建立的受控框架，使得人工智能系统的提供者或潜在提供者可以在监管监督下根据沙盒计划在有限时间内开发、训练、验证和测试（在适当情况下可在真实环境中进行测试）创新人工智能系统。”

人工智能监管沙盒可以实体、数字或混合形式建立，可以适用于实体和数字产品。

成员国有义务建立人工智能监管沙盒

成员国及其国家主管部门有义务建立人工智能监管沙盒（详见第8章）。每个成员国必须在2026 年 8 月 2 日之前建立至少一个人工智能监管沙盒，不过可以选择(i)在国家层面建立一个或多个人工智能监管沙盒，(ii)与一个或多个其他成员国的国家主管部门联合建立沙盒，或者(iii)加入现有沙盒。

建立人工智能监管沙盒的国家主管部门应当在适当情况下与其他相关国家主管部门合作，也可以将其他参与者纳入人工智能生态系统中。欧盟数据保护监督机构还可以为欧盟机构、团体、办事处和机关建立人工智能监管沙盒。

人工智能办公室将公布计划中和现有沙盒的清单。欧盟委员会还打算开发一个统一界面，其中包含与人工智能监管沙盒有关的信息，以便利益攸关方能够：

- 与人工智能监管沙盒互动；
- 向国家主管部门提问；和
- 寻求对创新人工智能产品、服务或商业模式的合规性的无约束力指导。

谁可以参与人工智能监管沙盒？

沙盒制度的目标群体是人工智能系统的提供者（或潜在提供者），但他们也可以与部署者及其他相关第三方合作提交申请。

沙盒制度中的一些具体规定旨在鼓励中小企业和初创企业参与，包括：

- 中小企业和初创企业使用沙盒通常应当免费；
- 对在欧盟设有注册办事处或分支机构的中小企业和初创企业提供优先准入；以及
- 中小企业和初创企业应当能够获得有关《人工智能法案》实施及其他增值服务的指导。

责任

对于因在沙盒中进行实验而给第三方造成的任何损害，参与人工智能监管沙盒的提供者和潜在提供者（包括中小企业和初创企业）仍要承担责任。但是，潜在提供者如有以下情形，不会被处以行政罚款：

- 遵守相关的沙盒计划以及参与条款和条件；并且
- （善意）遵守国家主管部门的任何指引。

沙盒制度的实施

为了避免欧盟内部出现分歧，欧盟委员会计划通过实施法案，来具体明确人工智能监管沙盒的建立、运行和监督，包括以下共同原则：

- 参与资格和遴选标准；
- 沙盒的申请、参与、监控、退出和终止程序；以及

- 参与者所适用的条款和条件。

这些实施法案旨在确保人工智能监管沙盒：

- 向任何符合公平透明的资格标准的提供者开放；
- 提供广泛平等的参与机会，并且满足参与的需求；
- 促进开发用于测试和说明与合规性学习有关的人工智能系统维度（例如准确性、稳健性和网络安全性）的工具和基础设施，以及降低对基本权利和整个社会造成风险的措施；
- 推动人工智能生态系统内相关参与者的参与（例如指定机构和标准化组织、测试和实验机构、研究和实验室以及欧洲数字创新中心），并且确保参与人工智能监管沙盒在整个欧盟得到统一认可（并具有相同的法律效力）。

国家主管部门的义务

国家主管部门必须：

- 配置足够资源，确保其沙盒制度符合《人工智能法案》的要求；
- 为沙盒参与者提供如何满足《人工智能法案》要求的指导；
- 为参与者出具合规退出报告书，详细说明沙盒中开展的活动、结果和学习成果，该报告随后可用于在符合性评估程序或相关市场监管活动中证明遵守了《人工智能法案》的要求；
- 向人工智能办公室和董事会提供年度报告（详见第8章），确定最佳实践、事件和经验教训。

国家主管部门将保留与沙盒活动有关的监督权，包括在必要时暂停或终止沙盒内开展的活动，以应对涉及基本权利或健康和安全的重大风险。

在沙盒内处理个人数据

为其他目的合法收集的个人数据可以在人工智能监管沙盒中使用，但须遵守《人工智能法案》规定的各项条件（必须满足所有条件才能进行获准的相关处理活动）。一些关键条件包括：

- 沙盒中部署的相关人工智能系统必须旨在维护重大公共利益（例如公共健康、能源可持续性、关键基础设施安全）；
- 个人数据的使用必须是必要的，并且不能用匿名或合成数据替代；
- 个人数据必须在独立且受保护的环境中处理，并且必须采取适当的技术和组织措施；以及
- 保留人工智能系统训练、测试和验证过程和基本原理的详细描述以及测试结果。

人工智能系统的真实世界测试

《人工智能法案》还允许在“真实世界条件”下对人工智能系统进行测试，但须符合特定条件。

> 《人工智能法案》对 *“真实世界条件下的测试”*的定义如下：
>
> *“在实验室或其他模拟环境之外的真实世界条件下对人工智能系统的预期用途进行临时测试，旨在收集可靠且可信的数据，评估和验证人工智能系统是否符合[《人工智能法案》]的要求”*。

只要符合《人工智能法案》的相关要求，真实世界测试就不算作是将相关人工智能系统投放市场或投入使用。（这些概念详见第 2 章）。

《人工智能法案》主要侧重于在人工智能监管沙盒之外对高风险人工智能系统进行真实世界测试。不过，《人工智能法案》也考虑到在国家主管部门的监督下，在人工智能监管沙盒框架内对人工智能系统（无论是否高风险）进行真实世界测试的可能性。

在上述两种情形下，真实世界测试都必须符合《人工智能法案》规定的各项条件（必须满足所有条件才能进行获准的测试，但在沙盒中进行测试的灵活性更高）。一些关键条件包括：

- 拟议的真实世界测试已取得相关市场监管机构的批准，并且已被录入欧盟高风险人工智能系统数据库中；
- 进行测试的提供者在欧盟境内设立（或已指定一名在欧盟境内设立的法律代表）；
- 测试时间最长6个月（可以再延长6个月，在沙盒环境中进行真实世界测试时，这一要求可以放

宽）；

- 真实世界测试的参与者受到适当的保护——必须取得这些参与者的知情同意，结果必须是可逆的（或可以被忽略），并且必须能够随时退出；以及

- 市场监管机构可以对真实世界测试的开展进行飞行检查。

提供者和潜在提供者将对其在真实世界测试过程中造成的任何损害承担责任。

参与沙盒和真实世界测试是否存在“挑选法域”的风险？

尽管《人工智能法案》旨在协调整个欧盟范围内与人工智能监管沙盒和真实世界测试有关的制度，但行业代表和利益攸关方无疑会密切关注相关动态，并且可以选择在被认为对行业最友好的司法管辖区参与沙盒和/或真实世界测试（包括如何确定参与沙盒或真实世界测试的相关责任）。

执法和治理

概览

- 《人工智能法案》建立了上市后监测、报告和信息共享流程。
- 高风险人工智能系统提供者是主要的义务承担者，必须建立上市后监测系统和程序，报告重大事件。
- 重大事件报告义务有时也适用于部署者。
- 报告的时间要求可能是即时的。
- 报告需要提交给事件发生地的成员国的市场监管机构；因此可能需要报告给多个有权部门。
- 执法方式多措并举：
  - — 欧洲数据保护监管局负责欧盟机构等。
  - — 欧盟委员会负责通用人工智能模型的提供者。
  - — 各成员国的主管部门负责其他相关事项。
- 根据违规情节的轻重，给予不同处罚。
- 受到影响的人有权获得每项决策作出原因的解释。

行动指南

高风险人工智能系统的提供者应当：

- 关注将于2026年2月2日通过的欧盟委员会上市后监测计划模板。
- 制定并实施上市后监测计划。
- 如果已经承担现有的上市后监测义务或者是受监管的金融服务提供商，请考虑是否可以将《人工智能法案》规定的义务整合到现有的上市后监测系统中。

高风险系统的提供者应当：

- 考虑是否已经承担其他同等义务；如果是，请检查是否有双重报告义务。
- 确保质量管理体系包括重大事件报告程序。
- 确保这些程序明确严重事件的性质（死亡、严重损害健康、侵犯基本权利等）以及事件是否具有广泛性。
- 确定汇报的对象

高风险系统的部署者应当：

- 制定预响应程序，以便在需要时进行报告。

提供者和开发者应当：

- 查看预计于2025 年 8 月 2 日发布的欧盟委员会指引。
- 持续关注相关内容，因为其将被重新评估。

非高风险人工智能系统的运营者应当：

- 确保遵守现有的所有产品安全法规。

行动指南

通用人工智能模型的提供者应当：

- 关注欧盟委员会有关执法安排的执行法案，并且考虑对征求意见进行反馈。

人工智能价值链中的所有组织都应当：

- 关注在成员国层面通过的执法规定，并且考虑对征求意见进行反馈。
- 注意，如有充分理由认为人工智能系统存在风险，要与市场监管机构合作。
- 注意，可能必须披露训练、验证和测试数据集以及源代码。

高风险系统的部署者应当：

- 确保他们能够清晰且有意义地说明人工智能的决策过程。

概述

《人工智能法案》为人工智能系统的事前要求以及事后监督执法的实施和管理提供了一个框架。事前要求已在前面的章节中进行了说明。本章主要涉及事后监督执行及治理结构。

执法机制针对两类风险：产品安全风险和基本权利风险。关于产品安全风险，《人工智能法案》以现有产品安全立法为基础，主要由国家市场监管机构执行。如果发现基本权利风险，市场监管机构应通知相关国家公共部门或机构，与其充分合作，保障基本权利。

与《人工智能法案》中基于风险分级的方法一致，该法规提供了多层次的执法架构，对具有不同风险的人工智能系统适用不同的机制。对于高风险的人工智能系统，《人工智能法案》首先规定了上市后监测义务，其次要求报告重大事件。重大事件报告义务有时也适用于部署者，因此部署者应当了解相关规则。

市场监管机构可以要求运营者采取一切适当措施，确保人工智能系统不会带来风险，必要时可要求将产品或人工智能系统从市场上撤回。对于违反《人工智能法案》规定的行为，还可处以巨额罚款。

对于通用人工智能模型，欧盟委员会拥有监督和执行《人工智能法案》中义务的专属权力。

《人工智能法案》中规定的治理架构包括在欧盟层面设立新的机构（人工智能办公室、欧洲人工智能委员会、咨询论坛和科学小组）和在国家层面设立新的机构（通知部门和市场监管机构），每个机构的职责和职能均有简要说明。这些机构之间的协调是有效实施和执行《人工智能法案》的关键所在。

本章讨论的主题包括：

- 上市后义务
- 市场监管机构
- 执法程序
- 基本权利保障机构
- 通用人工智能模型
- 处罚
- 对第三方的救济
- 治理

上市后义务

针对高风险人工智能系统的上市后监测系统

由于人工智能系统在上市后具有适应和继续学习的能力，因此在将人工智能系统投放市场后对其性能进行监测非常重要。序言155 条解释称，上市后监测系统旨在确保高风险人工智能系统的提供者能够考虑使用该系统的经验，从而确保系统持续合规和不断完善。

高风险人工智能系统的提供者在将系统投放市场之前编制技术文件时，必须包括上市后监测计划（第72(3)条和第11(1)条）。该计划必须符合欧盟委员会将截止于2026年2月2日通过的模板。上市后义务将确保可以识别出任何需要立即采取必要纠正或预防措施的情况（第3(25)条）。

第72条规定，上市后监测系统（以及系统文件）必须与人工智能技术的性质和系统风险相称。上市后监测系统必须在人工智能系统的整个生命周期内积极且系统地收集、记录和分析相关数据，以便提供者对持续合规情况进行评定。数据可以由部署者或其他人提供（但来自执法机构部署者的敏感运营数据除外）。在相关情况下，该系统还应包括与其他人工智能系统（包括设备和软件）交互的分析。

某些类型的高风险人工智能系统的提供者已有上市后监测系统的，可以将其根据《人工智能法案》承担的义务整合到现有系统中，但前提是这样做可以达到同等的保护水平。附件I第A部分中列出的欧盟统一立法所涵盖的高风险人工智能系统（即包括特定机械、玩具和医疗设备）符合这种情况。金融机构在市场上投放附件III第5点所列高风险人工智能系统（特别是信誉评估或者与人寿和健康保险有关的风险评估和定价）时，其内部治理、安排或流程也受欧盟金融服务法律规定的约束（第72(4)条）。

高风险人工智能系统的重大事件信息报告

高风险人工智能系统的提供者必须报告*“重大事件”*，提供者的质量管理系统必须包含与重大事件有关的程序（第17(1)(i)条）。通常，高风险人工智能系统的部署者必须向提供者报告重大事件。但是，如果部署者无法联系到提供者，则第73条规定的重大事件报告义务直接适用于部署者（第26(5)条）。因此，部署者也应当了解这些规定。欧盟委员会最迟将在2025年8月2日向提供者发布事件报告指南，并必须对其进行定期审查。

第3(49)条对重大事件进行了定义，是指人工智能系统的事件或故障直接或间接导致以下情况：

- 死亡或对人身健康造成严重伤害；
- 对关键基础设施的管理或运行造成重大且不可逆转的破坏；
- 违反保障基本权利的欧盟法律；或
- 对财产或环境造成重大损害。

重大事件必须在以下规定时限内报告。如有必要，提供者或部署者可以提交初步报告，再进行后续完善（第73(5)条）。

情况	时限
广泛侵权 或者 涉及关键基础设施的严重事件	立即 知悉这一事件后小于或等于2 天
人员死亡	知悉重大事件后小于或等于10天；或 在确定或怀疑重大事件与人工智能系统之间存在因果关系后立即（以较早者为准）。
其他情况（即严重损害健康、侵犯基本权利、严重损害财产或环境——除非这些情况达到广泛的程度）	知悉重大事件后小于或等于15天；或 在提供者确定人工智能系统与重大事件之间存在因果关系或合理可能性之后立即

报告后，提供者必须及时开展必要的调查，包括风险评估和纠正措施。在通知主管部门之前，提供者不得以任何可能影响对事件原因进行后续评估的方式改变人工智能系统。

严重事件报告必须向事件发生地的成员国的市场监管机构报告（第73(1)条）。如果严重事件影响多个成员国或影响多个行业，牵涉一个成员国内的多个市场监管机构，则需要提交多份报告。

市场监管机构必须在收到通知后的七天内采取适当措施（包括撤回或召回产品），还必须立即将任何重大事件通知欧盟委员会，无论其是否已采取行动（第73(8/11)条）。

非高风险人工智能系统

与非高风险产品有关的人工智能系统在投放市场或投入使用时必须是安全的。关于一般产品安全的(EU) 2023/988号条例和关于市场监管和产品合规的(EU) 2019/1020号条例适用于所有受《人工智能法案》管辖的人工智能系统，但上述两项条例为非高风险产品提供了安全保障（序言166和第74(1)条）。

(EU) 2019/1020号条例要求所有运营者在有理由认为存在第3(19)条规定的风险产品（定义见下文）时，应通知相关市场监管机构。《人工智能法案》在第3(19)条风险列表中增加了对人员基本权利的风险（第79(1)条）。

> “风险产品”是指可能对受到适用欧盟统一立法保护的一般人员健康和安全、工作场所健康和安全、消费者保护、环境、公共安全及其他公共利益造成不利影响的产品，其影响程度超出了与其预期用途有关或在所涉产品的正常或合理可预见的使用条件下（包括使用期限以及（如适用）投入使用、安装和维护要求）被认为是合理和可接受的范围。

市场监管机构

由于《人工智能法案》的执行通常需要当地机构的参与，因此成员国发挥着关键作用。各成员国必须指定至少一个市场监管机构，如果有多个机构，则必须指定其中一个机构为在成员国和欧盟层面与公众及其他相对人对接的单一联络点。成员国应将单一联络点通

知欧盟委员会，欧盟委员会将公布单一联络点名单（第153条和第70(1/2)条）。成员国必须在2025年8月2日之前遵守这些规定（第113(b)条）。

哪些实体将被指定为市场监管机构？

成员国指定市场监管机构具有一定的灵活性，可以新设一个机构专门执行《人工智能法案》，也可以将《人工智能法案》的要求整合到现有机构的框架中，该机构目前根据附件I第A部分所列的欧盟协调法律负责市场监管，或者监管受欧盟法律监管的金融或信贷机构（第74(3/6/7)条）。但是，对于生物识别、执法、移民、庇护和边境管控以及司法行政领域的高风险系统，成员国必须指定依据(EU) 2016/679号条例设立的国家数据保护机构或依据(EU) 2016/680号指令设立的监管机构（第74(8)条）。

如果人工智能系统涉及附件 I 第A部分所列欧盟统一立法已经涵盖的产品，并且这些法案规定的程序确保同等的保护水平和与《人工智能法案》相同的目标，则应适用行业程序，而不是第79条至第83条规定的国家层面执行程序（见下文“执法程序”部分）。

在这种情况下，不需要重复报告重大事件，提供者应根据其他法律进行报告（第73(9)条和第73(10)条）。这些例外情况特别适用于：

- 附件三类别的高风险人工智能系统，其提供者依据欧盟法律承担的报告义务与《人工智能法案》规定的义务相当。例如，网络安全法规中涉及的关键基础设施就有单独的事件报告义务，可被视为与《人工智能法案》规定的义务相当。然而，其他欧盟法律规定的报告义务是否被视为与《人工智能法案》规定的报告义务相当并不总是明确的；以及
- 高风险人工智能系统是设备的安全组件，或者本身就是设备，在关于医疗器械的(EU) 2017/745号条例和关于体外诊断医疗器械的(EU) 2017/746号条例中有相关规定。这两项法规都包含报告义务，要求如果重大事件导致(a)患者、用户或其他人死亡，(b)患者、用户或其他人的健康状况暂时或永久性严重恶化，或(c)严重的公共卫生威胁，则必须向主管部门报告。

但在上面两种情况下，如果侵权涉及侵犯基本权利，则仍必须根据《人工智能法案》进行通知，并且相关市场监管机构必须通知国家基本权利保障机构。

对于欧盟机构、机关、办事处和团体（欧盟法院行使司法职能除外）使用的人工智能系统，欧洲数据保护监管局将作为市场监管机构（第74(9)条）。

市场监管机构的权力

市场监管机构除拥有《人工智能法案》授予的其他权力外，还拥有(EU) 2019/1020号条例规定的所有广泛执法权。例如，该机构有权：

- 要求运营者披露与合规有关的文件、数据和信息。《人工智能法案》还规定，高风险人工智能系统的提供者可能会被要求披露：
  - 用于开发高风险人工智能系统的训练、验证和测试数据集，包括在适当且有安全保障的情况下，通过应用程序编程接口（API）或其他相关技术手段和工具实现远程访问（第74(12)条）；以及
  - 当基于提供者提供的数据和文件的测试或审计程序和验证已被用尽或者被证明不充分时，如果需要评估高风险人工智能系统是否符合第三章第二节规定的要求，则提供源代码（第74(13)条）；
- 进行不事先通知的现场检查和试购（第74(5)条）；
- 进行调查（当发现高风险人工智能系统对两个或多个成员国构成严重风险时，与欧盟委员会接洽）（第74(11)条）；
- 要求运营者采取适当措施，终止违规行为，包括形式违规（第83条）和消除风险（第79条至第82条）；
- 当运营者未采取纠正措施或违规情况持续存在时，采取适当措施，包括撤回或召回（第73(8)条、第79-83条）；
- 施加处罚（第99条至第101条）。

市场监管机构还应确保在真实世界条件下进行的测试符合《人工智能法案》（见第7章），其有权要求提供者或部署者修改测试或者暂停或终止测试（第76(3)条）。

保密信息的处理

市场监管机构获取的任何信息或文件均应按照第78条规定的保密义务处理。第78条的规定也适用于欧盟委员会、基本权利保障机构以及参与实施《人工智能法案》的自然人和法人。这些机构和人员在执行任务时，不仅应保护保密信息和商业秘密，还应保护

知识产权和源代码权利、公共和国家安全利益以及机密信息。

第78条的规定自2025年8月2日起适用。

执法程序

如前所述，如果已经存在提供同等保护水平且目标与《人工智能法案》相同的协调立法，则以下程序不适用。

存在风险的人工智能系统（第79条和第81条）

如果市场监管机构有充分理由认为人工智能系统存在风险（见上文定义），则必须对人工智能系统是否符合《人工智能法案》进行评估。

如有违规情形，市场监管机构应毫不延迟地通知并要求相关运营者采取一切适当的纠正措施，使人工智能系统合规或从市场上撤回或召回人工智能系统。市场监管机构应说明运营者必须满足合规要求的期限，但该期限不会超过 15 个工作日。

如果运营者在规定期限结束前未采取充分的纠正措施，市场监管机构应采取一切适当的临时措施，禁止或限制所涉人工智能系统进入其国内市场或投入使用，从其国内市场撤回或召回该产品或独立的人工智能系统。市场监管机构必须将其决定所依据的理由告知运营者。

如果违规行为不限于其国家领土范围内，则市场监督机构应毫不延迟地向欧盟委员会和其他成员国通报评估结果、其要求运营者采取的措施以及在运营者违规情况下采取的临时措施。

如果成员国市场监督机构或欧盟委员会在三个月内（若涉及违反第5条所述的禁止事项，则缩短至30天）未提出异议，则临时措施应视为合理。但是，如有异议提出，欧盟委员会应与市场监督机构和运营者协商，在六个月（如违反第 5 条，则为60天）内决定临时措施是否合理。如果合理，所有成员国应确保对所涉人工智能系统采取适当的限制措施，例如要求退出其市场。如果不合理，则将撤销临时措施。

上述规定不影响(EU) 2019/1020号条例第18条规定的运营者的程序权利，包括听证权。

提供者将人工智能系统归类为非高风险（第80条）

如果市场监管机构有充分理由认为提供者根据第6(3)条归类为非高风险的人工智能系统实际上是高风险的，则必须进行评估。

要遵循的程序与上面描述的非常相似，但第80条特别提到可以对相关提供者处以罚款。

市场监督机构在行使监督第80条实施的权力时，可以考虑欧盟高风险人工智能系统数据库中存储的信息（见下文“欧盟层面的治理：欧盟委员会的作用”部分）。

合规但存在风险的人工智能系统（第82条）

如果市场监管机构发现高风险的人工智能系统符合《人工智能法案》，但对人员的健康或安全、基本权利或者公共利益保护的其他方面构成风险，则应要求相关运营者采取一切适当措施，确保该风险不再存在。

形式违规（第83条）

市场监管机构发现例如CE标志应贴未贴、未指定授权代表或者未提供技术文件等情况的，市场监管机构应要求相关提供者在规定期限内纠正。

如果违规情况持续存在，市场监管机构应采取适当和适度措施，限制或禁止高风险人工智能系统投放市场，或者确保立即召回或从市场上撤回。

基本权利保障机构

除明确市场监管机构外，每个成员国还必须在2024年11月2日之前，指定负责监督和执行欧盟法中保护基本权利（包括反歧视权利）义务的公共机构或部门，这些义务与附录三中提到的高风险人工智能系统的使用相关，并将这些机构通知欧盟委员会。

市场监管机构发现基本权利受到威胁的，必须通知负责监督基本权利保障的相关国家公共机构。

上述机构有权要求并获取根据《人工智能法案》创建或维护的任何文件，只要获取这些文件对于有效履行其职责是必要的。相关公共机构或团体应将任何此类请求通知所涉成员国的市场监管机构，如果文件证明不充分，可请求市场监管机构运用技术手段组织对高风险人工智能系统进行测试（第77条）。

通用人工智能模型

欧盟委员会是负责监督和执行通用人工智能模型提供者义务的唯一机构。这是为了充分发挥欧盟层面的集中专业知识和协同效应（第88条）。但在实践中，人工智能办公室（见下文“治理”部分）将采取一切必要行动，监督《人工智能法案》关于通用

人工智能模型规定的有效实施，前提是欧盟委员会的组织权以及成员国和欧盟之间的职权划分不受影响。

人工智能办公室可以根据其监测活动的结果或市场监管机构的请求，主动调查通用人工智能模型提供者可能违规的行为。

对于基于通用人工智能模型的人工智能系统，在模型和系统由同一提供者开发的情况下，人工智能办公室具有市场监管机构的权力。

如果市场监管机构认为通用人工智能系统（可供部署者用于至少一个高风险用途）不符合《人工智能法案》，则市场监管机构必须配合人工智能办公室进行合规性评估。

当市场监管机构无法获取通用人工智能模型相关信息（从而无法完成对高风险系统的调查）时，市场监管机构可以请求人工智能办公室提供有关信息（第75条）。

处罚

任何违反《人工智能法案》的人——无论是自然人还是法人、公共机构还是欧盟或国家机构——都会因违规而受受到处罚。《人工智能法案》规定的处罚甚至超过了《通用数据保护条例》规定的处罚（最高可达20,000,000欧元或全球年营业额的4%）。最高罚款在整个立法过程中经过修改，最终定为35,000,000欧元或全球年营业额的7%。

罚款可由国家机构、欧洲数据保护监管局或欧盟委员会执行。欧洲数据保护监管局可对欧盟机构、机关和团体处以罚款。欧盟委员会可对通用人工智能模型提供者处以罚款。国家机构可对其他运营者处以罚款。

《人工智能法案》采取分级处罚方式，如下所示。

侵权理由	欧盟机构	所有其他人
	欧洲数据保护监管局施加处罚	国家机构当局施加处罚（除非是通用人工智能模型，由欧盟委员会施加处罚）。
		对于受制裁的企业，罚款上限为百分比金额或以下金额中的较高者。如果企业是中小企业，则罚款上限为较低者。对于其他受制裁的人，上限是规定的特定金额。
向通知机构或国家主管部门提供不正确、不完整或误导性的信息。	≤750,000欧元（第100(3)条）	≤上一年全球年营业总额的1%；或 ≤7,500,000 欧元（第99(5)条）
与高风险人工智能系统有关的义务。		≤上一年全球年营业总额的3%；或 ≤15,000,000欧元（高风险人工智能系统适用第99(4)条；通用人工智能模型适用第101(1)条）
与通用人工智能模型提供者有关的义务。		≤上一年全球年营业总额的 7%；或 ≤ 35,00,000 欧元（第99(3)条）
与禁止性行为有关的义务。	≤1,500,000欧元（第100(2)条）	≤上一年全球年营业总额的 7%；或 ≤ 35,00,000 欧元（第99(3)条）

令人疑惑的是，似乎对于未能履行第4条中关于人工智能素养义务的情况，并未设定任何处罚措施。

国家当局施加的处罚和罚款

成员国有责任制定有效、适度且具有劝诫性的处罚措施。这些措施可能包括罚款和非罚款措施或警告，必须在开

始实施之日前通知欧盟委员会（第99(1/2)条）。

处罚应根据具体情况进行。国家主管部门应当考虑具体情况的所有相关情节，并酌情考虑侵权行为及其后果的性质、严重程度和持续时间以及提供者的规模（第99(7)条）。

成员国层面的执法必须遵循适当的程序保障，包括有效的司法救济。

对欧盟机构、团体、办事处和机关的罚款

欧洲数据保护监督员有权对欧盟机构、机关和团体处以罚款。在作出罚款决定之前，欧洲数据保护监督员应将其初步调查结果告知欧盟机构，给予其陈述意见的机会。罚款不得影响该机构的有效运作，罚款所得资金应计入欧盟总预算。

对通用人工智能模型提供者的罚款

欧盟委员会可对侵权的通用人工智能模型提供者处以罚款（第101条）。与第十二章中自2025年8月2日起适用的其他处罚和罚款规定不同，第101条直到2026年8月2日才适用。

欧盟委员会将发布一项执行法案，详细说明执法的安排和程序保障。

在确定固定金额或定期罚款时，欧盟委员会应酌情考虑侵权行为的性质、严重程度和持续时间，并遵循比例原则和适当性原则。在作出罚款决定之前，欧盟委员会应将其初步调查结果告知通用人工智能模型的提供者，给予其陈述意见的机会。罚款的施加必须有适当的程序保障，包括欧盟法院的司法审查权。欧盟法院可以取消、减少或增加罚款金额。

对第三方的救济

向市场监管机构投诉（第85条）

欧盟和成员国法律已经为因使用人工智能系统而权利和自由受损的自然人和法人提供了一些有效的救济措施。尽管如此，《人工智能法案》引入了一种新的投诉机制，规定任何自然人或法人如果有理由认为存在违反《人工智能法案》的行为，都可以向主管市场监管机构进行投诉。

比较：根据《通用数据保护条例》，如果数据主体认为与其相关的个人数据的处理侵犯了受《通用数据保护条例》保护的权利，则数据主体有权向监管机构投诉涉嫌侵权行为。

相反，根据《人工智能法案》提出的投诉可能不仅涉及对投诉人权利的侵犯，还涉及《人工智能法案》的合规问题。此外，根据《通用数据保护条例》，只有数据主体才能申请救济，而根据《人工智能法案》，法人也可以提出投诉。

获取个案决策解释的权利（第86条）

根据《人工智能法案》，任何受到影响的人都有权要求部署者对高风险人工智能系统（关键基础设施系统除外）所做决定提供“明确且有意义的”解释。这些解释必须阐明相关决策程序和人工智能系统作出决定的主要因素（第86条）。

如有以下情况，可以行使上述权利：

- 部署者的决策主要基于高风险人工智能系统的输出；
- 这一决策对受到影响的人的健康、安全或基本权利造成不利影响，并产生具有法律影响或类似严重影响，。

比较：《人工智能法案》规定的解释权与《通用数据保护条例》规定控制者对自动化决策过程承担的义务相一致（《通用数据保护条例》第22条）。根据《通用数据保护条例》，控制者必须向数据主体提供有关自动化处理后果的逻辑和重要性的有意义信息。

《人工智能法案》第86条补充了《通用数据保护条例》规定的数据主体获取解释的权利；该条款更聚焦于人工智能，因为它要求部署者解释人工智能系统在决策中的作用。此外，《人工智能法案》还将这一权利赋予所有受到影响的人，包括法人。无论哪个机构有权执行《人工智能法案》第86条，就涉及个人数据处理的自动化决策而言，《通用数据保护条例》中规定的国家数据保护机构仍是负责执行控制者提供信息这一义务的主管机构。

对举报人的保护（第87条）

(EU) 2019/1937号指令关于保护举报欧盟法律违规行为人员的规定适用于举报违反《人工智能法案》的行为。

下游提供者的投诉（第89条）

《人工智能法案》允许下游提供者（通用人工智能系统的部署者）对可能违反该法规规定的行为进行投诉。

投诉可向人工智能办公室提出，且必须提供充分证

据。投诉内容至少应包括：

- 被投诉的通用人工智能模型提供者的详细信息及其联系人；
- 对相关事实的描述，以及所违反的条款；
- 投诉人认为存在侵权行为的理由；以及
- 发出请求的下游提供者认为相关的任何其他信息，包括酌情主动收集的信息。

下游提供者进行上述投诉的可能性使得人工智能办公室能够有效地监督《人工智能法案》的执行。

治理

已建立治理架构来协调和支持《人工智能法案》的实施，其目的是在欧盟和国家层面构建工作能力，整合利益攸关方，确保合作的可信性和建设性。

欧盟层面的治理：欧盟委员会的作用

根据《人工智能法案》的规定，欧盟委员会承担多项职责，包括制定和执行授权法案、制定和发布指引、制定标准和最佳实践，以及做出具有约束力的决定以有效实施《人工智能法案》。在实际操作中，这些职责将由人工智能办公室（隶属于通信网络、内容与技术总司的行政架构）承担，其职责是支持欧盟委员会的工作。

《人工智能法案》第八章规定了欧盟委员会与成员国合作必须执行的一项任务。欧盟委员会必须建立并维护一个欧盟数据库，覆盖第6(2)条中提到的高风险人工智能系统和根据第6(3)条不被视为高风险的人工智能系统。该数据库将包含：

- 由提供者或授权代表录入欧盟数据库的附录VIII的 A部分和B部分所列的数据；
- 由公共机构、机关或团体的部署者（或授权代表）录入欧盟数据库的附录VIII 的C部分所列的数据。

这些数据将向公众开放（执法、移民、庇护和边境管控领域的人工智能系统相关数据除外）。

《人工智能法案》设立的超国家机构

人工智能办公室的角色	行动
人工智能办公室由欧盟委员会经2024 年 1 月 24 日决定（C/2024/1459）设立。 人工智能办公室的职能是监督人工智能模型的发展，包括通用人工智能模型、与科学界的互动，在调查和测试、执法方面发挥关键作用，并践行全球使命（决定的序言第 5 条）。 人工智能办公室可聘请独立专家代表其进行评估。 人工智能办公室必须建立系统和程序来管理和防止潜在的利益冲突，并且必须发展联盟在人工智能领域的专业知识和能力。 人工智能办公室负责对通用人工智能系统的监督和控制（第75条）。	监测和执法：监测通用人工智能模型提供者的合规和义务履行情况。 调查：通过要求通用人工智能模型提供者提供文件和信息、进行评估和采取措施来调查侵权行为。 风险管理：在发现系统性风险时要求采取降低风险等适当措施，限制市场供应情况，、撤回或召回模型。 协调与支持：支持国家机构建立人工智能监管沙盒，促进合作与信息共享，鼓励和推动行为准则的制定。协调市场监管部门和欧盟委员会的联合调查。 咨询：就行为准则、实践准则和指引向欧盟委员会和欧洲人工智能委员会提出建议和书面意见。
欧洲人工智能委员会（委员会）的角色	**行动**
委员会由各成员国代表组成，负责就《人工智能法案》的一致和有效实施向欧盟委员会和成员国提供建议和协助。此外，委员会还发布指引和建议（第65条和第66条）。 代表任期为三年，可连任一次。代表可能是来自公共实体的具有人工智能专业知识和促进国家层面协调职权的个人。委员会主席由其代表之一担任。	协调与合作：国家主管部门与欧盟机构、团体、办事处和机关以及相关欧盟专家组和网络之间的协调与合作。 专业知识共享：收集并分享技术和监管专业知识、最佳实践和指导文件。 意见和建议：就《人工智能法案》的实施发表意见，特别是有关通用人工智能模型规则的执法，提出建议和书面意见（应欧盟委员会的要求或主动提出）。

欧洲人工智能委员会（委员会）的角色	行动
委员会必须设立两个专门的常设小组： • 通知机构常设小组：为与通知机构相关的问题提供合作与交流的平台 • 市场监管常设小组：作为《人工智能法案》的行政合作小组（ADCO）发挥作用。 委员会可根据需要设立其他常设或临时小组，以研究特定问题。 欧洲数据保护监督员和人工智能办公室以观察员身份出席委员会会议。其他国家和欧盟机构、团体或者咨询论坛的专家或代表可根据具体情况受邀出席。	协调：规范行政实践，促进共同标准的制定和统一理解的形成。 公众对人工智能的认知：致力于提升人工智能素养，提高公众对人工智能系统使用相关的益处、风险、保障措施、权利与义务的认知和理解。 国际合作：就人工智能的国际事务向欧盟委员会提供建议，并与第三国的主管机构及国际组织开展合作

咨询论坛的角色	行动
设立咨询论坛是为了确保利益攸关方参与《人工智能法案》的执行和实施（第67条）。 论坛成员由欧盟委员会任命，在利益攸关方代表的选择上保持平衡，涵盖行业主体、初创企业、中小企业、公民社会和在人工智能领域拥有公认专业知识的学术界。 论坛成员任期为两年，可延长至四年。两名联合主席从论坛成员中选出，任期为两年，可连任一次。 欧盟基本权利署（FRA）和欧盟网络安全局（ENISA）、欧洲标准化委员会（CEN）、欧洲电工标准化委员会（CENELEC）、欧洲电信标准协会（ETSI）为咨询论坛的常任成员。 咨询论坛可根据需要设立常设或临时小组，负责研究具体问题。 咨询论坛每年至少召开两次会议，并且可邀请专家和其他利益攸关方参加会议。	意见和技术专长：向欧洲人工智能委员会和欧盟委员会提供意见。应要求准备意见、建议和书面材料。 咨询小组：欧盟委员会在准备标准化请求或起草第41条所指的共同规范时，必须咨询论坛。 年度报告：编制并发布其活动的年度报告。

独立专家科学小组的角色	行动
成立科学小组是为了联合科学界支持欧盟委员会的执法行动（第68条）。 专家由欧盟委员会根据其当前在人工智能领域的科学或技术专长选出。 专家的人数由欧盟委员会与欧洲人工智能委员会协商后根据所需的专业知识需求确定，确保性别公平和在地域上具有代表性。 为了向科学小组提供履行任务所需的信息，应建立一种机制，允许科学小组请求欧盟委员会从提供者处获取文件或信息。 执行法案将明确科学小组及其成员如何发出警告并向人工智能办公室请求协助。	支持人工智能办公室在通用人工智能模型和系统的实施与执行方面的工作： • 向人工智能办公室警示可能存在的系统性风险。 • 开发用于评估能力的工具和方法。 • 就包括系统性风险在内的分类发表意见。 • 参与开发工具和模板。 • 支持市场监管机构：根据其请求，特别是在跨境市场监管活动方面提供支持。 • 协助进行第81条规定的欧盟保障程序。 根据需求支持成员国的执法活动： • 成员国可能需要为科学小组提供的建议和支持支付费用。 • 第 68(1) 条提到的执行法案将定义费用和可收回的成本。

国家层面的治理：国家主管部门

成员国在《人工智能法案》的实施和执法中发挥着至关重要的作用。为确保在欧盟内部和成员国之间有效实施、协调和配合，各成员国必须指定至少一个通知机构和一个市场监管机构。两者共同构成国家主管部门。对于欧盟机构、机关、办事处和团体使用的人工智能系统，欧洲数据保护监督员将作为监督机构。

通知机构的角色	行动
该机构负责建立和实施符合性评估机构的框架（第28条）。 该机构必须拥有足够数量的合格人员，这些人员具备信息技术、人工智能和法律等领域的必要专业知识，包括对基本权利的监督。 通知机构必须避免与符合性评估机构发生任何利益冲突，确保其活动的客观性和公正性。特别是，通知符合性评估机构的决定不得由对符合性评估机构进行评估的人员做出。	建立和执行程序：建立和执行符合性评估机构评估、指定、通知和监控的必要程序。与其他成员国的通知机构合作制定这些程序。 意见和指导：就《人工智能法案》的实施提供指导和意见，考虑欧洲人工智能委员会和欧盟委员会的意见，并咨询其他欧盟法律规定的国家主管部门（如适用）。 活动和服务限制： • 不得提议或提供符合性评估机构进行的任何活动。 • 不得以商业或竞争方式提供咨询服务。

市场监管机构的角色	行动
负责根据(EU) 2019/1020号条例（市场监管和产品合规）就市场监管和产品合规开展活动并采取措施。 各成员国将指定一个市场监管机构作为成员国和欧盟层面的公众及其他对应方的单一联络点。 欧洲数据保护监管局是《人工智能法案》项下欧盟机构、机关和团体的市场监管机构。 负责与生物识别有关的高风险人工智能系统的执法、移民、庇护、边境管控、司法和民主进程的高风险的市场监管机构应拥有强大的调查和纠正权限，包括访问所有个人数据和执行其任务所需的必要信息。 成员国必须促进市场监管机构和其他相关国家机构之间的协调。	除了上文描述的许多任务和职责外，市场监管机构还被赋予以下任务和职责： • 高风险人工智能系统的授权：成员国可以因公共安全、健康、环境保护或关键基础设施的特殊原因，临时授权特定的高风险人工智能系统在其领土内投放市场或投入使用，但需等待符合性评估（第46条）。 • 年度报告：向欧盟委员会和国家竞争管理机构报告监控活动和禁止实践，包括：（i）任何被识别出的可能与竞争法实施有关的信息；（ii）使用任何被禁止的行为；以及（iii）针对这些行为采取的措施。 • 意见和指引：就《人工智能法案》的实施提供指引和意见，考虑欧洲人工智能委员会和欧盟委员会的意见，并咨询其他欧盟法律规定的国家主管部门（如适用）。

我在哪里可以找到以下内容？

治理：第七章　　序言148-154、163和179

欧盟数据库：第八章　　序言131

《人工智能法案》：未来走向

概览

- 《人工智能法案》已于2024年8月1日生效。
- 大多数条款将于2026年8月2日起生效，其他条款将自生效之日起分阶段实施，过渡期为6至36个月。
- 欧盟委员会将制定授权法案和执行法案、指引、行为准则和标准。这些举措旨在提供与《人工智能法案》相关的实用指引、道德准则和技术规范，以确保该法规的有效实施。
- 欧盟委员会还于2024年7月向欧洲议会和理事会提交了更新版本的《人工智能责任指令》提案供审议。
- Bird & Bird的人工智能专家密切关注《人工智能法案》下即将出台的各项新举措，并协助您应对相关流程和合规要求。

行动指南

所有涉及人工智能系统的相关方都应积极关注本章中提到的立法和非立法举措的发展动态。

《人工智能法案》：未来走向

本章概述了《人工智能法案》的适用期限以及其法规下预计推出的相关举措。欧盟机构将《人工智能法案》视为一种新形式的“动态法规”，旨在通过二级立法及其他举措不断补充，以跟上技术发展的步伐。在未来几个月，《人工智能法案》预计将通过一系列授权法案、执行法案、指导文件、行为准则、实践准则和标准化要求。这些举措旨在为该法规提供实用指导、道德准则和技术规范，以确保其有效实施。

上述文件中的规定将极大地影响《人工智能法案》的有效实施和参与者履行其义务的能力。

因此，建议所有涉及人工智能系统的相关方积极关注欧盟委员会制定本章中提到的立法和非立法举措方面的工作。

Bird & Bird的监管和公共事务团队密切关注《人工智能法案》下即将出台的各项新举措，并协助您应对相关流程和合规要求。

《人工智能法案》的适用期限

继欧盟官方公报于2024年7月12日公布后，《人工智能法案》已于2024年8月1日生效。

相关适用日期如下。

2024年7月12日	《人工智能法案》在欧盟官方公报上公布，确定了该法规中具体条款的适用日期。
2025年2月2日	禁止性行为的规定适用（第二章）。 人工智能素养规则适用（第4条）。
2025年5月2日	通用人工智能的实践准则必须编制完成（第56(9)条）。
2025年8月2日	指定国家主管部门（第三章第4节）。 与通用人工智能（GPAI）相关的义务（第五章）。 治理（在欧盟和国家层面）（第七章）。 保密和处罚（不涉及通用人工智能的部分）（第十二章）。
2026年8月2日	《欧盟人工智能法案》的所有其他规定开始适用（除非以下另有较晚的适用日期）。
2027年8月2日	附件I中列出的高风险类别。 2025年8月2日之前投放市场的通用人工智能模型（第111条）。
2030年8月2日	在2026年8月2日之前已投放市场或投入使用，并拟由公共机构使用的高风险人工智能系统（不包括以下列出的系统）（第111条）。
2030年12月31日	在2027年8月2日之前已投放市场或投入使用的大规模 IT 系统（附件 X 所列）组件（第111条）。

6. 欧洲议会和理事会2024年6月13日(EU) 2024/1689号条例规定了有关人工智能的统一规则，并修订了(EC) 300/2008号条例、(EU) 167/2013号例、(EU) 168/2013号条例、(EU) 2018/858号条例、(EU) 2018/1139号条例和 (EU) 2019/2144 号例以及 2014/90/EU号指令、(EU) 2016/797号指令和(EU) 2020/1828号指令（《人工智能法案》）。文本具有欧洲经济区相关性，刊载于《欧盟官方公报》(OJ L)，2024年第1689号，2024年7月1日。

预计在2024年8月1日至2027年8月2日期间，欧盟委员会将通过各类文件来实施该法规。这些文件包括授权法案和执行法案、指导文件、行为准则、实践准则和标准化要求。除少数例外情况外，委员会尚未为这些举措的发布设定具体的截止日期。然而，可以推测，委员会将力争在相关条款的适用期限到来之前通过这些文件。

授权法案

若干条款将通过授权法案予以明确，由委员会负责制定，以明确义务和运作执行。第97条赋予欧盟委员会为期五年的授权法案制定权，该授权自2024年8月1日开始生效。欧盟委员会必须在授权期结束前九个月就这项授权提交报告。该授权期自动延长五年，除非欧洲议会或理事会在每个期限结束前三个月表示反对，该授权期限将自动延长五年。

如上所述，此类授权法案的通过并没有具体的截止日期。然而，可以推测，这些法案的通过将会早于《人工智能法案》中相关条款的适用期限（参见第113条）。

根据第97(4)条，在通过授权法案之前，欧盟委员会必须在准备工作期间进行公开咨询，并与相关专家组（由成员国专家组成）进行磋商。

一旦通过，欧盟委员会必须同时通知欧洲议会和理事会。只有当欧洲议会和理事会在通知后三个月（必要时可延长三个月）内均没有提出异议的情况下，授权法案才会生效。欧洲议会或理事会可以随时撤销这一权力，但这不会影响已有授权法案的有效性。根据2016年4月13日《关于完善立法的机构间协议》中确立的原则，欧盟委员会必须确保欧洲议会和理事会与成员国专家同时收到所有文件。此外，议会和理事会的专家应系统性地参与委员会专家组关于授权法案准备工作的会议。

《人工智能法案》预计在欧盟委员会认为必要的情况下通过以下授权法案：

- 第6(6/7)条： 如果有确凿且可靠的证据表明存在不应被纳入附件III或不应符合第6条第3款条件的人工智能系统，则可以通过以下方式修改第6条第3款：在第3款规定的条件基础上添加新的条件、修改现有条件或删除这些条件。
- 第7(1/3)条 ：修改附件III，增加、修改或删除高风险人工智能系统的应用场景；
- 第 11 (3) 条：在必要时修改附件IV，以确保根据技术进步，技术文件能够提供评估系统合规性所需的所有信息；
- 第43(5)条：根据技术进步，修改附件VI和附件VII并对其进行更新。
- 第45(6)条：修改第43(1/2)条，以使附件III第2至第8点所述的高风险人工智能系统需接受第三方合格性评估。
- 第 47(5)条：修改附件V，通过更新该附录中列出的欧盟合规声明的内容，以引入因技术进步而变得必要的要素。
- 第51(3)条：修改第51(1/2)条中列出的通用型人工智能模型的系统性阈值，并在必要时补充基准和指标，以适应技术发展的变化，例如算法改进或硬件效率提升，从而使这些阈值能够反映当前的技术水平。
- 第52(4)条：修改附件XIII，明确并更新系统性通用人工智能模型的标准；
- 第53(5)条：细化测量和计算方法，以便提供可比较和可验证的文件，从而促进对附件XI的合规情况；
- 第53(6)条：根据技术的发展，修改附件XI和附件XII 。

执行法案

《人工智能法案》第98(2)条赋予欧盟委员会根据182/2011号条例通过执行法案的权力[8]。执行法案旨在为特定立法法案的执行创造统一的条件（如有必要）。由成员国专家组成的“*执行*”委员会将协助欧盟委员会起草执行法案。

与授权法案的情况类似，预期实施法案的通过时间在文本中并未明确规定，唯有第72(3)条中提到的预期执行法案，其截止日期为2026年2月2日。因此，可以推测，相关执行法案将在《人工智能法案》中相关条款的适用期限之前通过（参见上述内容及第113条）。

在欧盟委员会认为必要的情况下，《人工智能法案》预计将通过以下执行法案：

- 第37(2)条：当成员国未能采取必要的纠正措施时，暂停、限制或撤销对通知机构的指定；
- 第41(1/4/6)条：与第67条所述的“咨询论坛”协商后，就高风险人工智能系统的要求或第五章第2节和第3节中规定的通用人工智能模型义务制定共同

规范。当涵盖本章第三节第2部分中相同要求的协调标准的参考文献在官方公报中发布时，委员会应废除第41（1）条中提到的执行法案。如果某成员国认为某一共同规范未完全满足本章第三节第2部分中规定的要求，委员会应评估该信息，并在适当情况下修改第41条第1款中提到的执行法案。

- 第50(7)条：按照第56(6)条规定的程序，批准为促进有效履行关于检测和标记人工生成或操纵内容义务而制定的行为准则。如果行为准则不够充分，欧盟委员会可以通过一项执行法案，制定一套共同规则，以落实第50条中针对某些人工智能系统提供者和部署者的透明度义务。

- 第56(6)条：批准适用于通用人工智能模型的行为准则，并使其在欧盟范围内具有普遍效力。如果在2025年8月2日之前无法最终确定行为准则，或者人工智能办公室认为其不够充分，欧盟委员会可通过执行法案制定共同规则，以落实第53条和第55条中规定的义务，包括第56(1)条中列出的问题。

- 第58(1)条：明确人工智能监管沙盒的建立、开发、执行、运行和监管的具体安排；

- 第60(1)条：明确高风险人工智能系统提供者的真实世界测试计划的具体要素；

- 第68(1)条：制定关于设立独立专家科学小组（“科学小组”）的规定，以支持《人工智能法案》的执行活动。

- 第72(3)条：不迟于2026年2月2日发布一项执行法案，规定详细条款，提供高风险人工智能系统提供商的上市后监测计划模板以及该计划中应包含的要素清单；

- 第92(6)条：规定人工智能办公室对通用人工智能模型进行评估的具体安排和条件，包括聘请独立专家的具体安排及其遴选程序；和

- 第101(6)条：针对通用人工智能模型提供者可能面临的罚款，制定详细的安排和程序保障。

欧盟委员会指引

“欧盟委员会指引”是由欧盟委员会服务机构编制的解释性文件，旨在为如何适用《人工智能法案》的具体规定提供实用和非正式的指导。

《人工智能法案》预计将通过以下欧盟委员会指引：

- 第6(5)条：在与欧洲人工智能委员会协商后，且不迟于2026年2月2日，具体说明第6条的实际执行情况，包括高风险和非高风险人工智能系统应用场景的实际示例的综合清单；

- 第63(1)条：考虑到微型企业的需求，以简化方式遵守质量管理体系的要素，但不得影响保护水平或高风险人工智能系统要求的合规性需要（此指南没有设定具体截止日期）；

- 第73(7)条：为促进重大事件报告义务的合规性提供指导。该指引必须不迟于2025年8月2日通过，并且必须由欧盟委员会定期评估；

- 第96条：关于本法案的实际执行的指引。虽然没有明确的截止日期。但相关条款自2026年8月2日起适用。特别的，欧盟委员会应就以下方面制定指引：——第8至15条和第25条中提到的要求和义务的适用；

  — 第5条所提及的被禁止的做法；

  — 有关重大修改条款的实际执行情况；

  — 第50条规定的透明度义务的实际执行；——有关《人工智能法案》与附件I所列的欧盟统一立法及其他相关欧盟法律之间关系的详细信息，包括在执法方面的一致性；以及

  — 第3(1)条规定的人工智能系统定义的适用。

行为准则和实践准则

行为准则

行为准则是具有自愿性质的文件，旨在为特定条件下人工智能的开发和使用制定道德准则和原则。这些准则和原则还旨在促进组织内部人工智能政策的制定，以自愿方式履行《人工智能法案》中的特定义务。

《人工智能法案》要求通过以下行为准则：

- 序言20和第4条：自愿行为准则旨在提高与人工智能开发、运营和使用相关人员的人工智能素养。

  — 虽然没有设定为提高人工智能素养而制定自愿实践准则的最后期限，但第4条中有关人工智能素养的相关规定将从2025年2月2日起适用。

- 序言165和第95条：行为准则旨在促进某些或全部适用于高风险人工智能系统的强制性要求在人工

智能系统中的自愿应用。这些准则根据系统的预期用途和所涉及的较低风险进行调整，并考虑到可用的技术解决方案和行业最佳实践，例如模型和数据卡：

— 为确保自愿行为准则的有效性，应当以明确的目标和关键业绩指标为基础，以衡量这些目标的完成情况；

— 准则的制定应以包容性方式进行，适当时应包括相关利益相关者的参与，例如企业、公民社会组织、学术界、研究机构、工会和消费者保护组织；以及

— 虽然未对旨在促进将适用于高风险人工智能系统的某些或全部强制性要求应用于人工智能系统的自愿行为准则的制定设定具体截止日期，但第95条中的相关规定将从2026年2月2日起适用。到2028年8月2日以及此后每三年，欧盟委员会将对此类自愿实践准则的影响和有效性进行评估。

实践准则

实践准则是正确遵守《人工智能法案》具体义务的核心工具。特别是，其中一项实践准则将详细说明《人工智能法案》中对通用人工智能模型和具有系统性风险的通用人工智能模型的提供者的规则。另一项实践准则将侧重于对人工生成或操纵的内容的检测和标记。组织应当能够依靠实践准则来证明其遵守了相关义务，这被称为“符合性推定”。

具体而言，《人工智能法案》要求欧盟委员会人工智能办公室推动与所有相关利益攸关方共同制定以下实践准则：

- 第50(7)条：欧盟层面的实践准则，旨在促进有效履行第50(2/4)条中关于检测和标记人工生成或操纵内容的义务。欧盟委员会可通过执行法案批准这些实践准则。虽然没有为制定自愿实践准则以促进有效履行第50(2/4)条中义务一事设定最后期限，但第50条中包含的相关规定将从2026年2月2日起适用。

- 第56(1/3)条：到2025年5月2日，制定通用人工智能模型的实践准则。这些准则将充分考虑国际做法以及各种不同观点，与相关国家主管部门合作，并在适当情况下与公民组织及其他相关利益攸关方和专家协商，其中包括根据《人工智能法案》设立的独立专家“科学小组”。

到2028年8月2日及此后每三年，欧盟委员会必须评估自愿实践准则的影响和有效性。

2024年7月30日，欧洲人工智能办公室发布了关于参与起草首个通用人工智能实践准则的意向征集公告。有意者可最晚在2024年8月25日表达参与意向。据欧盟委员会称，该行为准则将通过一个迭代的起草过程完成，目标是在《人工智能法案》于2024年8月1日生效后的九个月内，即2025年4月完成起草。该实践准则将促进《人工智能法案》中针对通用人工智能模型规则的正确实施。

根据第56(6)条，欧盟委员会可决定批准上述实践准则，并通过执行法案赋予该实践准则在欧盟范围内的普遍效力。如果该实践准则被认为不够充分，欧盟委员会将为相关义务的履行制定共同规则。

此外，2024年7月30日，人工智能办公室根据《人工智能法案》启动了关于可信赖的通用人工智能模型的咨询，具体涉及用于训练通用人工智能模型的内容摘要模板及随附指南。提交意见的截止日期为2024年9月10日。

标准

初步标准化工作

在《人工智能法案》通过之前，支持其实施的欧洲标准起草工作就已启动。欧盟委员会于2023年5月22日通过了关于人工智能统一规则的提案，并以委员会执行决定C(2023)3215的形式正式采纳。

上述执行决定要求欧洲标准化委员会（CEN）和欧洲电工标准化委员会（CENELEC）在不迟于2025年4月30日起草以下关于人工智能的新欧洲标准或欧洲标准化可交付成果：

- 关于人工智能系统风险管理体系的欧洲标准和/或欧洲标准化可交付成果；

- 关于用于构建人工智能系统的数据集的治理和质量的欧洲标准和/或欧洲标准化可交付成果；

- 通过人工智能系统的日志记录功能进行记录保存的欧洲标准和/或欧洲标准化可交付成果；

- 关于人工智能系统用户透明度和信息提供的欧洲标准和/或欧洲标准化可交付成果；

- 关于人类监督人工智能系统的欧洲标准和/或欧洲标准化可交付成果；

- 关于人工智能系统准确性规范的欧洲标准和/或欧洲标准化可交付成果；

- 关于人工智能系统稳定性规范的欧洲标准和/或欧

洲标准化可交付成果；

- 关于人工智能系统网络安全规范的欧洲标准和/或欧洲标准化可交付成果；
- 关于人工智能系统提供者的质量管理体系（包括上市后监测流程）的欧洲标准和/或欧洲标准化可交付成果；以及
- 关于人工智能系统合格评定的欧洲标准和/或欧洲标准化可交付成果。

向欧洲标准化委员会（ CEN）和欧洲电工标准化委员会（CENELEC）提出的这项标准化请求是基于欧盟委员会2022 年“欧洲标准化年度欧盟工作计划”第63号行动，其目标是确保人工智能系统的安全可信。

为了起草这些标准，欧洲标准化委员会（CEN）和欧洲电工标准化委员会（CENELEC）成立了一个名为“ CEN-CENELEC JTC 21 人工智能”的专门联合技术委员会。欧洲标准化委员会（CEN）和欧洲电工标准化委员会（CENELEC）还与信息和通信领域的独立非营利性标准化组织欧洲电信标准协会（ETSI）合作起草该标准。

《人工智能法案》标准化请求

《人工智能法案》第40(2)条要求欧盟委员会在该法规生效后毫不拖延地提出统一欧盟人工智能标准的标准化请求，内容涵盖：

- 《人工智能法案》第三章第2节规定的所有要求；以及
- 如适用，涵盖《人工智能法案》第五章第2节和第3节中规定的义务的标准化请求。

上述请求修改了欧盟委员会执行决定C(2023)3215中包含的请求，这也体现在欧盟委员会在2024年2月发布的2024 年标准化工作计划中。事实上，工作计划的第15项行动要求“修改标准化请求以支持欧盟的人工智能政策”，从而要求根据《人工智能法案》的最终文本修改欧盟委员会的决定。

根据《人工智能法案》第40(2)条，标准化请求还应要求提供报告和文档流程方面的可交付成果，以提高人工智能系统的资源性能。此类请求可能包括减少高风险人工智能系统在其生命周期内对能源和其他资源的消耗，以及通用人工智能模型的节能开发。欧盟委员会应在咨询欧洲人工智能委员会和相关利益攸关方（包括根据《人工智能法案》设立的利益攸关方咨询论坛）后起草申请。

此外，欧盟委员会在向相关欧洲标准化组织发出标准化请求时，应说明标准必须是明确且一致的。这一前置要求包括为附件I中列出的现有欧盟协调立法所涵盖的各个领域的产品制定的标准。这些标准旨在确保在欧盟投放市场或投入使用的高风险人工智能系统或通用人工智能模型符合《人工智能法案》规定的相关要求或义务。

委员会需在2028年8月2日之前以及此后每四年提交一份报告，审查关于通用人工智能模型节能开发的标准化交付成果的进展情况。在此背景下，委员会还需评估是否需要采取进一步措施或行动，包括具有约束力的措施或行动。该报告必须提交给欧洲议会和理事会，并向公众公开。

责任

欧盟委员会修订提案以与《人工智能法案》保持一致

最后值得注意的是，2024年7月底，欧盟委员会向欧洲议会和理事会提交了其针对人工智能调整非合同民事责任规则的提案更新版本（即《人工智能责任指令》（ AI Liability Directive，简称AILD））。该提案最初由委员会于2022年9月提出，旨在通过一套规则解决人工智能特定用途所产生的风险，重点关注对基本权利和安全的尊重。目前的修改旨在使《人工智能责任指令》提案与已完成的《人工智能法案》保持一致。

值得注意的是，新提案对第4条进行了修订，增加了部署人工智能系统的公司可能承担的责任。如果这些部署者“未监控人工智能系统的运行，或在适当情况下中止其使用”，或者未使用“具有足够代表性的”输入数据，将被推定对所造成的损害负责。

欧洲议会负责起草这份文件的首席起草人（“报告员”）、德国基督教民主党议员阿克塞尔·沃斯（Axel Voss）此前曾请求欧洲议会研究服务中心进行“替代影响评估”，以评估在《人工智能法案》正式通过后，人工智能责任指令是否仍然有必要。尽管拟议的人工智能责任指令的未来仍不确定，但可能会以简化形式继续实施。

人工智能指南撰稿人

作为一间市场领先的技术型律师事务所，我们在12个司法管辖区被知名法律评级机构Legal 500评为人工智能第一等级律所（欧洲法律指南中的首个同类排名）和电信、媒体和技术（TMT）第一等级律所，被钱伯斯指南评为全球跨司法管辖区电信、媒体和技术第一等级律所。我们通过深入理解人工智能技术开发和部署所涉及的技术复杂性而脱颖而出。这些专业知识使我们能够有效地与开发人员和商业团队合作，使用他们的语言，从一开始就提出正确的问题。我们的国际人工智能团队由120多名专家组成，几乎涵盖了人工智能这一变革性技术同法律法规的所有交集。从处理开创性的知识产权诉讼和指导客户应对复杂的监管变化，到实施有效的治理框架和创新商业和合同安排。

如果您对内容有任何疑问，请联系以下任何一位撰稿人或您的Bird & Bird联系人。您还可以在我们的AI Hub 中了解有关最新的人工智能发展动态。

比利时

Benoit Van Asbroeck
合伙人
+3222826067
benoit.van.asbroeck@twobirds.com

Francine Cunningham
监管和公共事务总监
+3222826056
francine.cunningham@twobirds.com

Paolo Sasdelli
监管和公共事务顾问
+3222826076
paolo.sasdelli@twobirds.com

中国

姚捷
合伙人
+862123121218
christine.yiu@twobirds.com

龚钰
法务总监
+861059335699
james.gong@twobirds.com

屈尘
律师
+861059335568
harry.qu@twobirds.com

芬兰

任熙
律师，鸿鹄罗杰联营团队
+862123121269
emma.ren@twobirdslawjay.com

Tobias Bräutigam
合伙人
+358962266758
tobias.brautigam@twobirds.com

法国

Anne-Sophie Lampe
合伙人
+33142686333
anne-sophie.lampe@twobirds.com

Cathie-Rosalie Joly
合伙人

+33142686742
cathie-rosalie.joly@twobirds.com

Cen Zhang
高级律师

+33142686000
cen.zhang@twobirds.com

Delphine Frye
高级律师

+33142686054
delphine.frye@twobirds.com

Dr. Miriam Ballhausen
合伙人

+4940460636000
miriam.ballhausen@twobirds.com

Dr. Nils Lölfing
法律顾问

+4921120056000
nils.loelfing@twobirds.com

Oliver Belitz
法律顾问

+4969742226000
oliver.belitz@twobirds.com

Dr. Simon Hembt
高级律师

+4969742226000
simon.hembt@twobirds.com

意大利

Gian Marco Rinaldi
法律顾问

+390230356071
gianmarco.rinaldi@twobirds.com

波兰

Aleksandra Cywinska
高级律师

+48225837875
aleksandra.cywinska@twobirds.com

Aleksandra Mizerska
律师

+48225837900
aleksandra.mizerska@twobirds.com

Andrzej Stelmachowski
律师

+48225837977
andrzej.stelmachowski
@twobirds.com

Izabela Kowalczuk-Pakula
合伙人

+48225837932
izabela.kowalczuk-pakula@
twobirds.com

Dr. Maria Jurek
高级律师

+48225837839
maria.jurek@twobirds.com

Marta Kwiatkowska-Cylke
法律顾问

+48225837964
marta.kwiatkowska-cylke@twobirds.com

Pawel Lipski
合伙人

+48225837991
pawel.lipski@twobirds.com

Tomasz Zalewski
合伙人

+48225837946
tomasz.zalewski@twobirds.com

西班牙

Joaquín Muñoz
合伙人

+34917906007
joaquin.munoz@twobirds.com

Feyo Sickinghe
法律顾问

+31703538904
feyo.sickinghe@twobirds.com

英国

Shima Abbady
高级律师

+31703538984
shima.abbady@twobirds.com

Alex Jameson
高级律师

+442078507139
alex.jameson@twobirds.com

Ian Edwards
合伙人

+442079056377
ian.edwards@twobirds.com

Katerina Tassi
高级律师

+442074156066
katerina.tassi@twobirds.com

Katharine Stephens
合伙人

+442074156104
katharine.stephens@twobirds.com

Liz McAuliffe
律师

+442074156787
liz.mcauliffe@twobirds.com

Nora Santalu
律师

+442079826513
nora.santalu@twobirds.com

Ruth Boardman
合伙人

+442074156018
ruth.boardman@twobirds.com

Toby Bond
合伙人

+442074156718
toby.bond@twobirds.com

Will Bryson
高级律师

+442074156746
will.bryson@twobirds.com

twobirds.com

Bird & Bird