

KPMG

毕马威

# 人工智能 (AI) 国际金融监管初探



毕马威金融行业研究中心

2024年09月

44,1665

# 执行摘要

随着经济社会的发展与科学技术的进步，人工智能已经被广泛应用于工作生活的各个领域。各国纷纷将人工智能纳入国家战略，期待其为经济发展提供新的增长点与动能。特别是2022年Chat-GPT的问世，向世界展现了人工智能巨大的发展潜力与应用前景。但同时，人工智能技术也可能带来系统性歧视、对数据和隐私权的侵犯以及信息茧房等问题。为推动技术以人为本，有效应对潜在风险，各国开始立足于自身的技术优势与资源禀赋，采用不同的人工智能治理与监管路径。

在吸收国际组织研究成果的基础上，欧盟率先推出人工智能伦理规范、法案与框架公约等，利用治理格局优势地位积极推动国际监管，并借助欧盟统一市场的力量，通过产业链向全球施加影响。英国采取了比较灵活的监管路径，并利用地缘优势与脱欧后的自由空间，优先促进技术的创新与发展，吸引科

技公司并促进产业转化。以美国、中国、新加坡为代表的具有人工智能技术优势的国家，则在国家规划的整体引领下布局促进技术发展与产业培植，兼顾发展中可能产生的风险与问题，积极推动人工智能的治理与监管。

本文立足于全景扫描，在收集研读人工智能相关文献的基础上，对国际组织、欧盟、英国、美国、中国等具有代表性主体的监管体系进行了分析，并对具有里程碑意义的文件及内容要点进行了整理与解读。同时，我们亦结合各国家/地区金融领域的人工智能监管动态进行了总结、分析与展望，并对金融机构人工智能技术应用提出了合规行动建议。



# 目录

## 03 人工智能监管概览

## 05 主要国家/地区人工智能监管

- 05 一、国际组织
  - 1.1 国际组织体系概览
  - 1.2 重要监管文件分析
  - 1.3 金融领域的人工智能监管与影响
- 10 二、欧盟/欧洲
  - 2.1 欧盟/欧洲体系概览
  - 2.2 重要监管文件分析
  - 2.3 金融领域的人工智能监管与影响
- 15 三、英国
  - 3.1 英国体系概览
  - 3.2 重要监管文件分析
  - 3.3 金融领域的人工智能监管与影响
- 21 四、美国
  - 4.1 美国体系概览
  - 4.2 重要监管文件分析
  - 4.3 金融领域的人工智能监管与影响
- 27 五、中国
  - 5.1 中国体系概览
  - 5.2 重要监管文件分析
  - 5.3 金融领域的人工智能监管与影响

## 36 结语

# 人工智能监管概览

## 国际人工智能监管体系概况

根据毕马威收集的全球人工智能监管相关文献分析，国际上人工智能立法较为活跃的国家与地区有欧盟、美国、英国、中国、新加坡以及香港等。从已发布的文件来看，各区域/国家普遍采取了促进创新，以人为本，以风险为导向的原则来实施或推进人工智能的监管，具有强制约束力的监管规则较少，仅占18%（且一半为美国州级的立法），多数采取了原则、伦理、指南、讨论文件及报告等更为软性的形式（图1）。

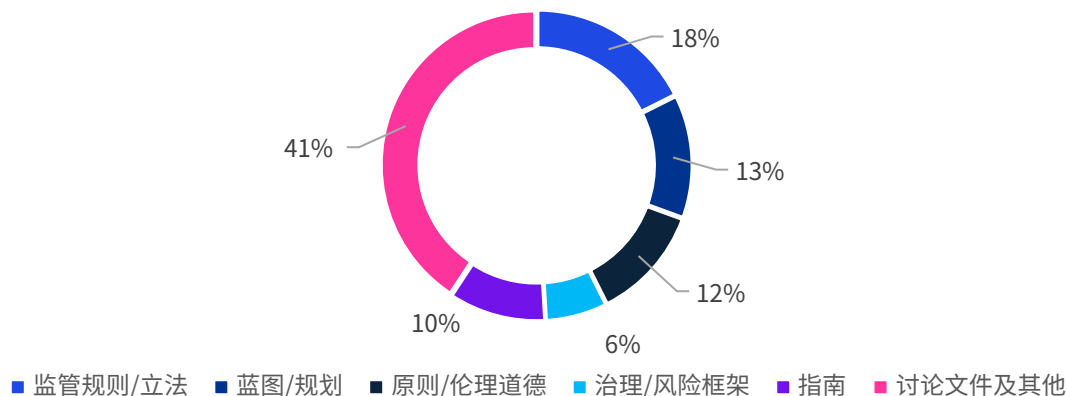
从各区域/国家的人工智能立法特点来看，国际组织积极推出人工智能原则、伦理规范，并通过宣言、倡议、联合声明等方式推动全球整体的人工智能治理与伦理约束。欧盟的人工智能监管体系更为完整，全面覆盖规划、原则、伦理、框架、指引、监管规则等各种类型；欧盟同时积极参与国际社会人工智能监管规则的制定，其推出的《欧盟人工智能法案》为全球首部人工智能领域的全面监管

法规，《人工智能与人权、民主和法治框架公约》为全球首个关于负责任地使用人工智能的具有法律约束力的国际条约，具有较强的国际影响力。

英国的人工智能监管体系非常有特点，它整体采用了一种政府引领，以部门为主导、有利于创新的做法来实施监管。它采用了相对软性的方式，不主张仓促立法与推出强制规范，文件多以软性的规划、框架、指南等形式颁布，同时英国监管机构还实施了大量研究并与社会各界人士广泛探讨，以保持对人工智能实践与风险的敏锐度。

美国的人工智能监管立法具有州领跑联邦政府的特点，一批州级的法律规范已率先落地，联邦部门的立法在政府执行令的统一引导下逐步推进，但尚未正式推出具有强制遵循意义的法律规范。

图1：国际人工智能监管相关文献分布情况



数据来源：毕马威自行收集整理。数据收集期间为2017年7月至2024年7月底

# 人工智能监管概览

中国较早就开始实施人工智能规划及治理活动，虽然已出台的文件数量不多，但种类较为丰富，已涵盖发展规划、监管规则、治理原则、伦理规范、倡议与宣言等。同时，中国也积极参与国际社会的人工智能监管治理，推动人工智能产业发展并构建自身的国际影响力。

从金融领域来看，目前金融监管机构对于人工智能监管的整体立法并不十分活跃，各区域/国家金融监管部门主要根据整体性的人工

智能监管政策与要求，在各自的职能领域针对人工智能应用进行规范。但随着上位法及人工智能专门法规的陆续出台，相关金融监管机构亦在酝酿人工智能领域的金融监管要求。

以下我们将选择国际组织、欧盟、英国、美国等四个人工智能监管领域具有代表性及领先意义的主体开展分析。我们亦对中国人工智能监管的进展与动态进行了总结，后续将单独撰写与国际实践比较分析的专题文章。



# 主要国家/地区人工智能监管

## 一、国际组织

### 1.1 国际组织体系概览

国际组织较早就开始了关于人工智能的探讨，从其目前已发布的文件看，均为不具有强制约束力的原则、伦理建议、倡议宣言等，旨在推动相关方以人为本，以负责任的方式设计、开发和应用人工智能。相关内容为不同地区、国家的人工智能立法提供了有力的支持与丰富的借鉴，包括标准定义、原则导向、伦理规范、工具方法等。

2019年，世界经济合作组织（OECD）与二十国集团（G20）先后推出了《人工智能原则》，但原则并不具有约束力。2021年，联合国教科文组织（UNESCO）发布《人工智能伦理问题建议书》，对人工智能涉及的伦理问题进行了深入探讨，并提供了评估工具与政府行动建议。2023年，七国集团（G7）领导人发布《就广岛人工智能进程的联合声明》，并配套推出了指导原则与行为准则。2023年11月，首届全球人工智能（AI）安全

峰会上，包括中国、美国、欧盟在内的 28 个国家，签署了《布莱切利 AI 宣言》，这也是全球第一份针对人工智能这一快速新兴技术的国际性声明（图2）。

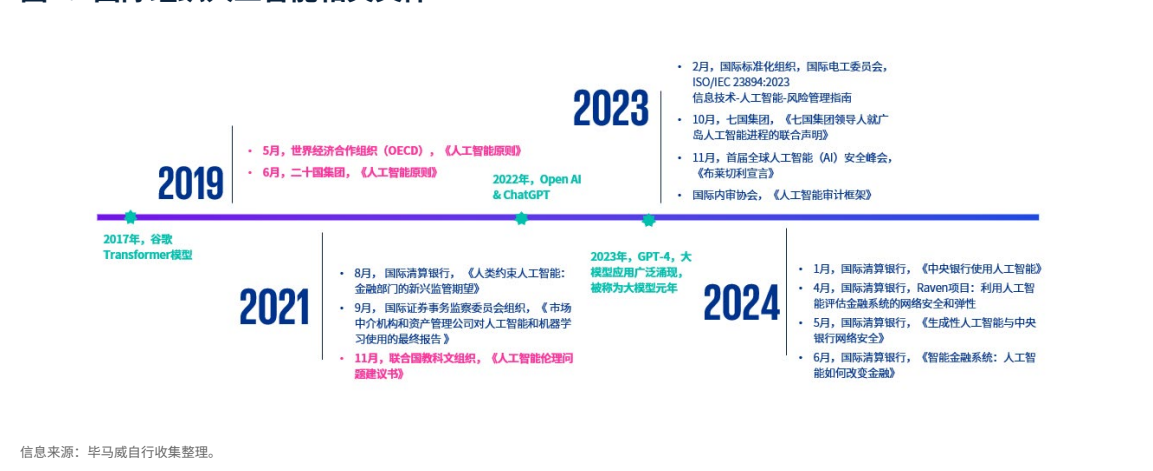
### 1.2 重要监管文件分析

以下选择三份重要的国际组织人工智能监管文件进行分析。

#### OECD《人工智能原则》和G20《人工智能原则》（2019年）

在全球层面，从国际组织发布的人工智能相关文件来看，较为重要且被广泛引用的两份文件是世界经济合作组织（OECD）于2019年5月发布的《人工智能原则》（首个人工智能政府间标准），以及二十国集团（G20）于2019年6月发布的《人工智能原则》，后者借鉴了前者的提议<sup>1</sup>（图3）。在后续各地区、国家、组织发布的人工智能监管文件中，关

图2：国际组织人工智能相关文件



<sup>1</sup> 部分内容引用自FSI Insight on policy implementation No 35, Humans keeping AI in check-emerging regulatory expectations in the financial sector.

# 主要国家/地区人工智能监管

图3: OECD和G20提出的人工智能原则

世界经济合作组织 (OECD)	二十国集团 (G20)
包容性增长、可持续发展和福祉 Inclusive growth, sustainable development and well-being	包容性增长、可持续发展和福祉 Inclusive growth, sustainable development and well-being
人权和民主价值观, 包括公平和隐私 Human rights and democratic values, including fairness and privacy	以人为本的价值观与公平 Human-centered values and fairness
透明度和可解释性 Transparency and explainability	透明度和可解释性 Transparency and explainability
稳健性、保障性和安全性 Robustness, security and safety	稳健性、保障性和安全性 Robustness, security and safety
责任制 Accountability	责任制 Accountability

信息来源: 毕马威根据OECD《人工智能原则》和G20《人工智能原则》整理。

于人工智能的标准定义、原则等, 均吸收利用了两份文件的成果。

## 《人工智能伦理问题建议书》(2021年11月)

### ——首个全球级的人工智能伦理标准

2021年11月, 联合国教科文组织 (UNESCO) 采纳了一套关于人工智能伦理的建议《人工智能伦理问题建议书》(以下简称建议书), 由193个成员国一致通过, 这是首个全球级的人工智能伦理标准。它建议会员国在自愿基础上适用建议书的各项规定, 特别是根据各自国家的宪法实践和治理结构采取适当步骤, 包括必要的立法或其他措施, 依照包括国际人权法在内的国际法, 使建议书的原则和规范在本国管辖范围内生效<sup>2</sup>。《建议书》详细讨论了其目标、价值观、原则以及政策领域

的行动(图4), 并提供了两项核心工具, 准备状态评估 (readiness assessment methodology) 与伦理影响评估 (ethical impact assessment)。

准备状态评估工具 (RAM) 包括一系列定量和定性问题, 旨在收集与一个国家的人工智能生态系统相关的不同方面的信息, 包括法律和监管、社会和文化、经济、科学和教育以及技术和基础设施方面。其中每个维度都有一系列定量和定性评估问题, 使RAM有别于其他现有的准备状态评估工具。其问题设置主要覆盖六大维度, 包括通用维度、法律维度及其8个细分维度、社会与文化维度及其5个细分维度、科学与教育维度及其10个细分维度、经济维度及其3个细分维度, 以及技术与基础设施维度及其4个细分维度。其中法律

<sup>2</sup> 部分内容引用自UNESCO, 《人工智能伦理问题建议书》中文版。



## 主要国家/地区人工智能监管

维度包含人工智能政策与监管细分维度（图5），由于与本文研究主题高度相关，详细列示以供参考。

伦理影响评估工具（EIA）考虑了设计、开发和部署人工智能系统的整个过程，包含一系列问题以及评估表格等，可以用于人工智能系统向公众发布之前和之后对其风险及影响

进行评估。伦理影响评估工具最初设计目的是帮助参与人工智能系统采购的公共部门官员，为他们提供一系列问题，以确保他们购买的人工智能系统符合《建议书》中规定的伦理标准。然而，它也可以被开发人员或私营部门和其他地方的其他人使用，以促进人工智能系统符合道德的设计、开发和部署。

图4：UNESCO《人工智能伦理问题建议书》提出的价值观与原则



信息来源：毕马威自根据UNESCO《人工智能伦理问题建议书》整理。





# 主要国家/地区人工智能监管

图5：准备状态评估工具人工智能政策与监管评估指标

AI政策与监管——评估指标	
<b>定性</b>	<p>1、贵国是否制定了国家人工智能战略？如果没有，您是否有任何对人工智能监管产生间接影响的立法或战略（例如，数据隐私或反歧视法，或数字战略）？</p>
	<p><b>评估</b></p> <p>1.1、国家人工智能战略的有效性是否得到了评估？</p> <p>1.2、国家人工智能战略或同等战略是否包括伦理成分？</p> <p>1.3、国家人工智能战略或同等战略是由一个多元化的团队（包括男性和女性，少数民族等）提出？</p> <p>1.4、国家人工智能战略或同等战略是否是在与不同利益相关者协商后制定的（学者、企业高管、民间社会等）？</p> <p>1.5、国家人工智能战略或同等战略是否包括人工智能对人权的影响？</p> <p>1.6、人工智能战略或同等战略是否包括详细的实施计划？</p> <p>1.7、是否有专门的机构/工作组负责领导人工智能战略的实施或同等工作？</p> <p>1.8、人工智能战略或同等战略是否包括为其建议的措施分配预算？</p> <p>1.9、人工智能策略或同等策略是否要求在某些领域部署之前进行人工智能影响评估？</p> <p>1.10、人工智能战略或同等战略是否包括监测和评估规定？</p>
	<p>2、贵国是否颁布了任何具有约束力的人工智能法规或软性法律（例如采购指南）？如果贵国尚未制定任何人工智能法规，是否正在制定此类法规？</p>
	<p><b>评估</b></p> <p>2.1、这种具有约束力的人工智能监管的有效性是否得到了评估？</p>

信息来源：毕马威自根据UNESCO《人工智能伦理问题建议书》整理。

## 《七国集团领导人就广岛人工智能进程的联合声明》（2023年10月）

2023年10月30日，七国集团（G7）领导人就“广岛人工智能进程”发表联合声明，强调要开发先进的人工智能系统，充分利用生成式人工智能带来的机遇和潜力，打造以人类为中心的包容的人工智能治理框架。基于联合声明，七国集团领导人当日还发布了《广岛进程先进人工智能系统开发组织国际指导原则》和《广岛进程先进人工智能系统开发

组织国际行为准则》<sup>3</sup>。文件提出了11项与风险相匹配的原则：

- 在先进人工智能系统的整个开发过程中，包括在部署和投放市场之前，采取适当的措施，以识别、评估和减轻整个人工智能生命周期的风险；
- 在部署及投放市场后，识别和减轻漏洞，并在适当的情况下识别滥用事件和模式；

<sup>3</sup> 内容引用自中国国际科技交流中心，科技动态 | 七国集团领导人就“广岛人工智能进程”发表联合声明。

# 主要国家/地区人工智能监管

- 公开报告先进人工智能系统的能力、局限性、适当和不适当使用的领域，以支持确保足够的透明度，从而有助于加强问责制；
- 致力于在开发先进人工智能系统的组织之间，包括与行业、政府、民间社会和学术界，实现负责任的信息共享和事件报告；
- 制定、实施和披露基于风险的人工智能治理和风险管理政策，包括隐私政策和缓解措施，特别是针对开发先进人工智能系统的组织；
- 投资并实施强有力的安全控制，包括整个人工智能生命周期的物理安全、网络安全和内部威胁防护；
- 在技术可行的情况下，开发和部署可靠的内容认证和来源机制，如水印或其他技术，使用户能够识别人工智能生成的内容；
- 优先考虑研究以减轻社会、安全和保障风险，并优先考虑对有效缓解措施的投资；
- 优先发展先进的人工智能系统，以应对世界上最大的挑战，特别是但不限于气候危机、全球卫生和教育；
- 推动制定并酌情采用国际技术标准；
- 实施适当的数据输入措施，保护个人数据和知识产权。

## 1.3 金融领域的人工智能监管与影响

从金融领域来看，国际清算银行（BIS）、国际证券事务监察委员会组织（IOSCO）等国际组织一直在积极推进金融领域人工智能相关的研究，自2021年起陆续发布金融领域的人工智能研究报告，研究主题覆盖市场中介机构和资产管理公司对人工智能和机器学习的应用，人工智能在中央银行的应用，中央银行的生成式人工智能与网络安全，人工智能如何重塑金融业等，其研究内容主要聚焦于人工智能在金融业的运用现状，机会与挑战，以及蕴含风险的探讨，并未推出具有遵循意义或操作指导的治理原则、风险框架、指引等规范。国际清算银行金融稳定研究所（BIS-FSI）的研究报告也提出，建议区分面客（customer-facing）、非面客（Non customer-facing）两类场景搭建不同的监管框架与措施。展望未来，随着人工智能技术的持续发展与金融领域应用场景的不断扩大，预计未来将有更多国际组织积极作为，促进金融领域人工智能监管或引导规范的出台。

# 主要国家/地区人工智能监管

## 二、欧盟/欧洲

### 2.1 欧盟/欧洲体系概览

欧盟自2016年起就不断探索对人工智能技术应用的监管体系构建，其以欧洲整体层面为视角，致力于构建协调一致、适合欧洲发展的人工智能治理体系。欧盟监管的政策重点是，以基本人权为底层逻辑，侧重于保障个人权利。以数据安全为阵地延伸，通过人工智能使用数据与数据应用场景划分其风险等级。其治理模式是与风险相匹配的分层治理。监管机制是采取分级监管、多种监管，并实行高额的行政处罚，且将公私机构一起纳入监管范围<sup>4</sup>。从国际竞争的视角来看，欧盟一方面利用自身国际影响力，积极拟定并输出伦理指南、全面监管法案、国际公约等，发挥制度优势影响国际层面的人工智能治理与监管。另一方面以欧盟统一市场为基础，延伸向域外乃至全球供应链相关方施加影响力。

欧洲的人工智能监管体系较为完善，已颁布的文件虽然数量不多，但其内容全面覆盖规划蓝图、监管立法、原则伦理、治理框架、

指引指南等类型。每份文件正式出台前，均经历了多轮的草拟议案、意见征询、讨论反馈与修订等（图6）。

### 2.2 重要监管文件分析

以下选择三份欧盟人工智能领域监管具有里程碑意义的文件进行重点分析。

#### 《值得信赖的人工智能伦理指南》（2019年4月）

——不具备强制约束力的伦理指南，提出了值得信赖的人工智能框架

基于2018年12月份的初稿与意见征询，2019年4月8日，欧盟委员会下设的人工智能高级别专家组正式发布了《值得信赖的人工智能伦理指南》，提出了实现值得信赖的人工智能全生命周期框架（图7）。

值得信赖的人工智能包含三要素，符合法律的（lawful），合乎道德的（ethical）和

图6：欧盟/欧洲人工智能相关文件



<sup>4</sup> 部分内容引用自《人工智能监管制度概览》，国家社会科学基金重点项目“人工智能研发与应用风险治理的财税法协同机制研究”（立项编号：21AFX021）的阶段成果，来源人民法院报。



# 主要国家/地区人工智能监管

图7：值得信赖的人工智能框架指南



数据来源：毕马威根据《值得信赖的人工智能伦理指南》翻译整理。

具有稳健性的 (robust)。开发、部署和应用人工智能需符合四项伦理原则，包括：尊重人的自主性 (respect for human autonomy)、预防伤害 (prevention of harm)、公平 (fairness) 和具有可解释性 (explicability)。为达成以上目标，框架还提出了七项关键要求，包括：人的能动性 & 监督 (human agency and oversight)，技术稳健性与安全性 (technical robustness and safety)，隐私与数据治理 (privacy and data governance)，透明性 (transparency)，多样性、非歧视性与公平性 (diversity, non-discrimination and fairness)，环境和社会福利 (environmental and societal well-being)，责任制 (accountability)。

## 《欧盟人工智能法案》 (2024年5月)

——全球首部人工智能领域的全面监管法规

《欧盟人工智能法案》出台经历了漫长的酝酿过程，除欧盟内部的立法程序与反复商讨外，还曾多次遭遇空客、西门子等大型企业，以及谷歌、微软、OpenAI等科技巨头的公开反对与抵制。2021年4月，欧盟委员会提出了一项《旨在协调人工智能规则的法规提案》。2023年6月，欧洲议会通过《人工智能法案》谈判授权草案。12月，欧洲议会、欧洲理事会和欧盟委员会达成临时协议，将把《人工智能法案》制定为法律。2024年3月13日，

# 主要国家/地区人工智能监管

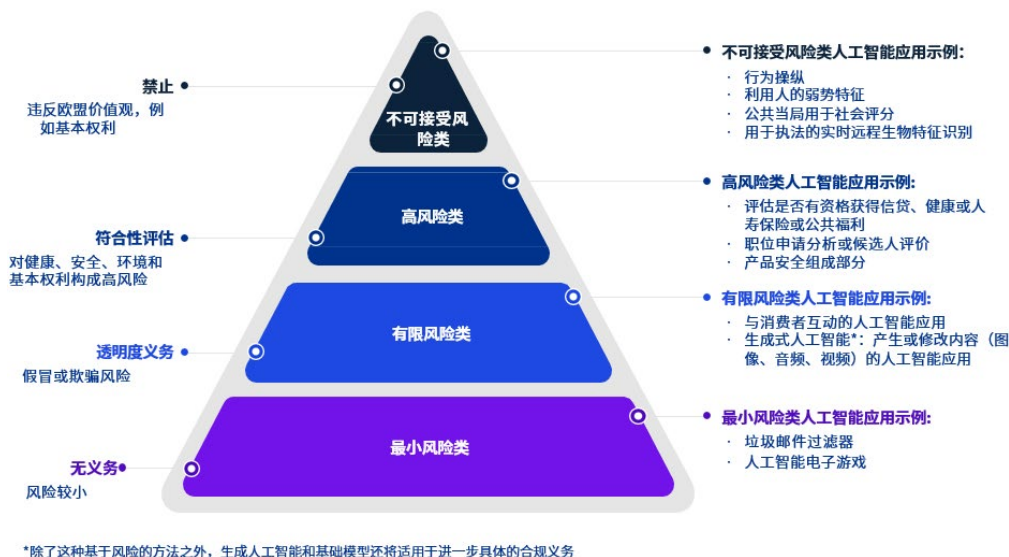
欧洲议会批准《人工智能法案》。5月21日，欧洲理事会正式批准欧盟《人工智能法案》（以下简称“法案”），这是第一个具有约束力的全球人工智能横向监管法规，为欧盟人工智能系统的使用和供应设定了一个共同的法律框架，具有标志性意义。

《法案》是一份全面的文件，它旨在提供人工智能的明确定义，使欧盟范围内关于人工智能的法律法规与其他欧盟法律法规保持一致。《法案》的主要目标是建立一个统一和横向的法律框架，促进人工智能系统的使用，同时提供高水平的保护，防止其不利影响。

从主要内容来看，《法案》共13章，包括总则，禁止的人工智能实践，高风险人工智能系统，特定人工智能系统的供应者与使用者需遵循的透明度义务，通用人工智能模型，支持创新的措施，治理，欧盟高风险人工智

能数据库，上市后监测、信息共享与市场监督，行为准则与指引，授权与委员会程序，处罚，以及最后条款。《法案》的重点是保护安全和基本权利，它根据“基于风险的方法”，为不同的人工智能应用系统风险水平，引入了分级的监管要求系统（图8）。一些存在“不可接受”风险的人工智能系统被禁止。一些可能对人们的健康、安全或基本权利产生不利影响的“高风险”人工智能系统得到授权后可以使用，但需要遵守一系列进入欧盟市场的要求和义务。由于缺乏透明度而造成“有限风险”的人工智能系统将受到信息和透明度要求的约束，而存在“最小风险”的人工智能将不受进一步义务的约束。此外，《法案》还为通用人工智能（GPAI）模型提供了具体规则，考虑到具有“高影响能力”的通用人工智能模型可能会构成系统性风险并对内部市场产生重大影响，《法案》还对其提出了更严格的要求。

图8：欧盟《人工智能法案》的人工智能系统分级监管



数据来源：根据毕马威国际Decoding the EU AI Act翻译整理。

## 主要国家/地区人工智能监管

从适用范围来看，《法案》为整个人工智能产业链上的相关方提供了一个合规义务框架，旨在约束与欧盟市场有连接点的所有主体，包括 供应商（providers）、进口商（importers）、分销商（distributors）、产品制造商（product manufacturers）、部署商（deployers）和用户（users）等。

从《法案》生效时间与处罚后果来看，法案中规定禁止的做法将在生效6个月后适用；通用型人工智能（GPAI）的相关义务和规则将在生效12个月后适用；法案生效24个月后，该法案完全适用，但一些高风险AI系统的规则将在生效后36个月适用。不遵守法案的后果可能包括限制市场准入和根据违法程度处以巨额罚款。罚款可能从750万欧元或全球营业额的1%，到3,500万欧元或全球营业额的7%不等，具体取决于人工智能系统的类型、公司的规模和违法行为的严重程度。

### 《人工智能与人权、民主和法治框架公约》 (2024年5月)

——全球首个关于负责任地使用人工智能的具有法律约束力的国际条约

《人工智能与人权、民主和法治框架公约》的起草工作于2019年启动，当时人工智能特设委员会（CAHAI）的任务是审查这种工具的可行性。在其授权之后，人工智能委员会（CAI）于2022年继任，起草并谈判了该文件。2024年5月18日，欧洲理事会在斯特拉斯堡正式通过了具有历史意义的《人工智能与人权、民主和法治框架公约》（以下简称“框架公约”），这标志着全球首个关于负责任地使用人工智能的具有法律约束力的国际条约正式诞生。它旨在确保人工智能系统生命周期内的活动完全符合人权、民主和法治，

同时有利于技术进步和创新。《框架公约》未对技术进行规范，本质上是技术中立的。

《框架公约》共8章，36条，主要内容包括：总则、总体义务、与人工智能系统生命周期内的活动相关的原则、补救措施、风险和不利影响的评估和缓释、公约的执行、后续机制与合作、以及最后条款。

从各国需履行的《框架公约》责任来看，主要包括：

#### 基本原则

- 人的尊严和个人自主（Human dignity and individual autonomy）；
- 平等和不歧视（Equality and non-discrimination）；
- 尊重隐私和个人数据保护（Respect for privacy and personal data protection）；
- 透明度和监督（Transparency and oversight）；
- 问责制和负责任（Accountability and responsibility）；
- 可靠性（Reliability）；
- 安全创新（Safe innovation）。

#### 补救措施、程序性权利和保障措施

- 记录有关人工智能系统及其使用的相关信息，受影响人员可获得相关信息；
- 信息必须足以使相关人员质疑通过使用该系统或在很大程度上基于该系统做出的决定，并质疑该系统本身的使用；



# 主要国家/地区人工智能监管

- 有效保障向主管当局提出申诉的可能性；
- 在人工智能系统对人权和基本自由的享受产生重大影响的情况下，就人工智能系统的应用向受影响的人提供有效的程序保障、防护设施和权利；
- 提供通知，说明一个人正在与人工智能系统交互，而不是与人类交互。

## 风险与影响的管理要求

- 以持续迭代 (iterative) 的方式，对人权、民主和法治的实际和潜在影响进行风险和影响评估；
- 通过实施这些评估，制定充分的预防和缓释措施；
- 保留当局对人工智能系统的某些应用实施禁令或暂停的可能性（“红线”）。

从《框架公约》的适用范围和约束效力来看，涵盖公共当局 (public authorities)，亦包括代表其行事的私立行为者，以及私营部门及个体 (private actors) 对人工智能系统的使用。《框架公约》为缔约方在管理私营部门时遵守其原则和义务提供了两种方式：缔约方可以选择直接接受《框架公约》相关条款的约束，或者作为一种选择，采取其他措施遵守条约的规定，同时充分尊重其在人权、民主和法治方面的国际义务。此外，《框架公约》缔约方不必将条约的规定适用于与保护其国家安全利益有关的活动，但必须确保这些活动尊重国际法和民主体制和进程。

《框架公约》不适用于国防事务，也不适用于研发活动，除非人工智能系统的测试可能干扰人权、民主或法治。

## 2.3 金融领域的人工智能监管与影响

除通用人工智能监管政策外，从金融领域人工智能监管政策与要求来看，欧洲银行管理局、欧洲保险与职业养老金局等已发布的监管规则，以及德国、法国、卢森堡、荷兰等国家发布的适用于本国的治理框架、总体原则等，基本与欧盟通用人工智能监管要求保持一致。此外，欧盟金融监管机构与各国监管机构还发布了多项人工智能相关的专题研究报告。针对《法案》与《框架公约》后金融领域的人工智能监管职责来说，欧盟国家可以将这项任务分配给其人工智能监管机构或现有的国家金融监管机构。与此同时，作为欧洲重要金融监管机构，欧洲央行 (ECB) 在监管《法案》的要求方面没有设置角色，但将继续从审慎的角度审查信贷模型。但随着《法案》与《框架公约》的生效与实施，预计未来欧盟及各国也将调整或细化监管政策，进一步明确金融领域人工智能监管的职责分工与监管要求。

从金融领域的影响来看，依据《法案》的风险分级，部分金融行业的人工智能应用被划分为高风险，例如银行的人工智能信用评级系统被指定为高风险系统，因为这可能会对个人或群体造成不公平歧视。又如保险领域，为健康或人寿保险单定价的人工智能系统也归类为高风险。此类人工智能系统必须满足高标准的稳健性和准确性，必须在强大的风险管理框架内运行，并且必须确保人工监督和对其输出信息的正确理解。这些要求将适用于《法案》生效两年后部署的新系统，为金融机构履行合规义务保留了窗口期。而依据《法案》与《框架公约》的强制性要求，金融机构应用人工智能还需建设相应的治理架构、持续的风险评估与管理流程，并配备相应的岗位与专家等。

# 主要国家/地区人工智能监管

## 三、英国

### 3.1 英国体系概览

英国在人工智能监管路径设计方面主张平衡创新与发展<sup>5</sup>，整体采用了一种政府引领，以部门为主导、有利于创新的做法来实施人工智能监管。政府部门、监管机构与特设机构通过发布政策白皮书、讨论文件、研究简报、政策反馈、政府应答、实施效果评估等方式，持续通过深入的研究、广泛的探讨与动态的论证，并充分征询社会公众与商业企业的信息输入与反馈意见，促进相关部门不断审视并检讨原则与目标的实现情况，用较低成本的方式达成整体人工智能治理与监管目标，并促进公众的理解与信任。此外，英国高度重视并积极参与国际标准的制定，并将推广促进有利于英国发展的国际标准作为降低合规成本，促进国家竞争力的有效手段。在最新发布的《科学技术框架》，英国将成为国际治理优胜者与领导者列入拟于2030年实现的战略目标。

从已发布的人工智能相关文件来看，英国采用的监管方法比较温和，不主张仓促推出立法与监管规范，文件多以软性的规划、框架、指南等形式颁布。同时，英国对人工智能的发展变化保持密切关注，发布了大量的研究报告、白皮书、讨论文件等，并将持续实施相关调研与研究，以保持对风险的敏锐感知与及时响应（图9）。

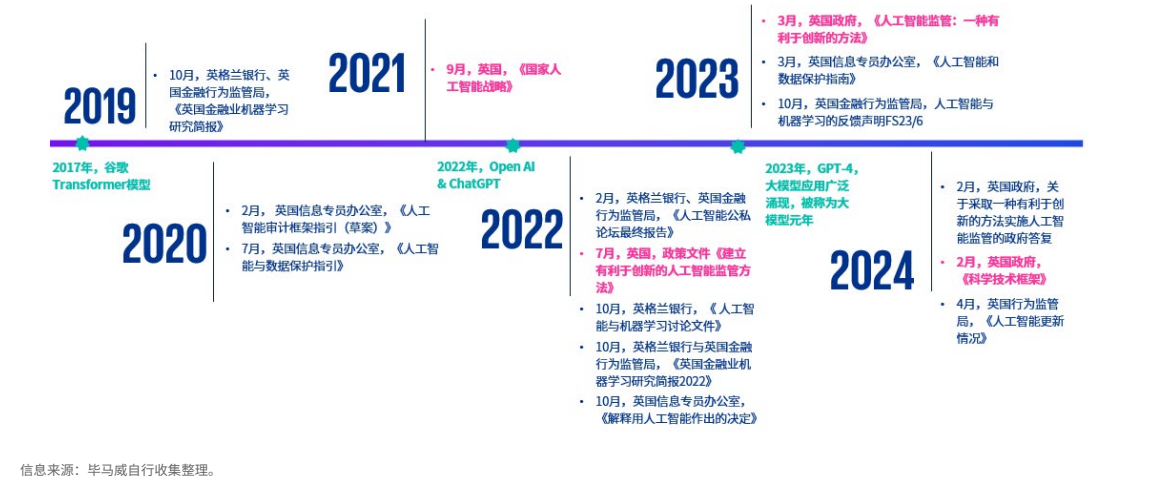
### 3.2 重要监管文件分析

以下将对英国四份重要的监管政策与文件进行分析。

#### 《国家人工智能战略》（2021年9月）

在2021年9月发布的《国家人工智能战略》（以下简称《战略》）中，英国政府承诺在治理和监管人工智能方面采取有利于创新的国家立场，提出了“使英国成为人工智能超级大国的十年计划”以及关于高效治理人工

图9：英国人工智能相关文件



<sup>5</sup> 引用自《人工智能监管制度概览》，国家社会科学基金重点项目“人工智能研发与应用风险治理的财税法协同机制研究”（立项编号：21AFX021）的阶段性成果，来源人民法院报。

## 主要国家/地区人工智能监管

智能的路径设想（图10）。《战略》还对后续人工智能即将采取的十项监管举措进行了部署说明。

### 《建立有利于创新的人工智能监管方法》 (2022年7月)

承接《战略》中的监管举措部署，英国政府在2022年7月的一份政策文件《建立有利于创新的人工智能监管方法》（以下简称《政策文件》）中概述了政府的立场。这份文件阐述了政府关于实施一个适度的、经得起未来考验的、有利于创新的人工智能监管框架的建议。为了促进创新并支持蓬勃发展的人工智能生态系统，英国采取的方法是：

- 情景化的 (Context-specific)：承认人工智能是一种动态的通用技术，由此产生的风险主要取决于其应用的情景；

- 支持创新和基于风险的 (Pro-innovation and risk-based)：要求监管机构专注于可能导致真实、可识别、不可接受的风险水平的人工智能应用，而不是试图对低风险或假设风险的人工智能应用实施控制，以避免扼杀创新；

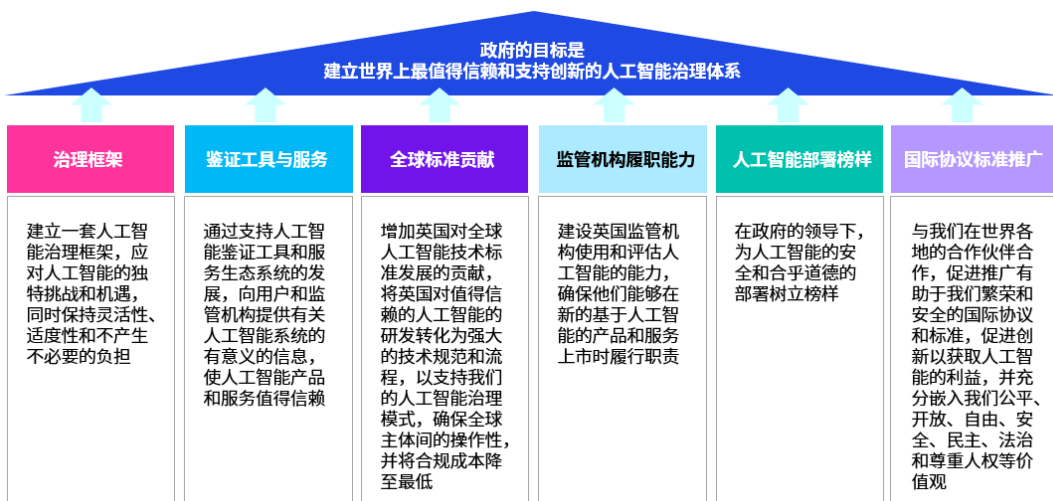
- 连贯的 (Coherent)：确保体系简单、清晰、可预测和稳定；

- 适度且适应性强的 (Proportionate and adaptable)：要求监管机构首先考虑更轻盈的选择，如指导或自愿措施等。

针对人工智能技术的特征，《政策文件》还建议制定一套跨部门原则，由监管机构将其发展为特定部门或领域的人工智能监管措施。6项跨部门原则包括：

- 确保人工智能安全使用；

图10：英国《国家人工智能战略》中的政府目标与实现路径



信息来源：毕马威根据《国家人工智能战略》整理



# 主要国家/地区人工智能监管

- 确保人工智能在技术上是安全的，并在功能上按设计运行；
- 确保人工智能具有适当的透明性和可解释性；
- 将公平考虑嵌入人工智能；
- 明确法人对人工智能治理的责任；
- 明确纠偏或质疑的途径。

## 《人工智能监管：一种有利于创新的方法》 (2023年3月)

2023年3月，在《政策文件》基础上，充分吸纳各方反馈意见与政策影响评估见解，英国正式发布了《促进创新的人工智能监管方法》白皮书（以下简称《监管方法》），总体而

言将通过一种基于原则，中央职能持续监测并促进协作，能够灵活适应的方法来实施监管（图11）。新监管框架以五项原则为基础，以指导和影响人工智能在所有经济部门的负责任开发和使用。这些原则将在非法定基础上发布（on a non-statutory basis），并由现有监管机构实施。

- 安全性、保障性和稳健性（Safety, security and robustness）；
- 适当的透明度和可解释性（Appropriate transparency and explainability）；
- 公平性（Fairness）；
- 责任制与治理（Accountability and governance）；

图11: 英国《促进创新的人工智能监管方法》中的人工智能监管策略



数据来源：毕马威根据《促进创新的人工智能监管方法》翻译整理。

# 主要国家/地区人工智能监管

- 争议与纠偏 (Contestability and redress)。

监管框架围绕四个关键要素构建，旨在增强现有监管机构的能力，促进整个监管格局的一致性。四大关键要素为：

- 根据人工智能的独特特性定义人工智能，以支持监管机构协调；
- 采用针对特定情景的方法；
- 提供一套跨部门原则，指导监管机构应对人工智能风险和机遇；
- 提供新的中央职能，支持监管机构实施人工智能监管框架，最大限度地提高迭代方法的效益，并确保框架的一致性。

在实施方面，为促进跨部门的横向协作，《监管方法》确定了一些必要的中央支持职能，以确保总体框架对风险做出适度但有效的反应，同时促进整个监管领域的创新：

- 监测和评估总体监管框架的有效性和原则的实施情况，包括实施在多大程度上支持了创新。这将使英国能够保持响应能力，并在必要时调整框架，包括在人工智能能力和最新技术发展的背景下需要调整以保持有效的地方；
- 评估和监测人工智能在整个经济中产生的风险；
- 进行全景扫描和差距分析，包括召集行业一同为应对新兴人工智能技术趋势提供一致的信息；
- 支持试验台和沙盒计划，帮助人工智能创新者将新技术推向市场；

- 提供教育和意识宣传，为企业提供清晰的指引，并赋予公民权力，让他们的声音被听到，作为框架不断迭代的组成部分；
- 促进与国际监管框架的互用性。

## 《科学技术框架》（2024年2月）

2024年2月，英国发布《科学技术框架》政策文件（以下简称“框架”），它制定了政府到2030年使英国成为科技超级大国的战略愿景。它将人工智能确定为五大关键技术之一，并指出监管在创造人工智能蓬勃发展的环境方面发挥着重要作用。

《框架》阐明了监管与规则方面的愿景，即英国始终处于制定技术标准和制定国际法规的前沿，并将充分利用脱欧后的自由。监管将有利于创新，刺激对科技的需求，吸引投资，同时代表英国的价值观并保护公民。政府将利用其科学技术优势和国际关系，来确保对法规和技术标准的影响力。并希望能在2030年达成以下目标：

- 一个有利于创新、易于引导并促进广泛商业科学和技术应用的法规和标准体系。监管机构对每一项关键技术的作用都将是明确的，包括它们在哪些领域跨部门和监管职责。监管机构将有权支持英国创新者的创新测试成本，使其具有国际竞争力；
- 相对于其他国家，迅速采取行动，为关键技术制定规则，并在适当的情况下制定法规，以增加这些领域创新者的确定性。英国将在制定国内法规时考虑国际背景，以加强英国的先发优势。涵盖关键技术的法规将处于世界领先地位；

# 主要国家/地区人工智能监管

- 领导制定关键技术标准和法规的国际行动。英国将在世贸组织、七国集团、二十国集团、经合组织、北约、欧洲理事会、英联邦和联合国发挥积极作用。英国将成为关键技术国际规则和公约的召集人和顶级作者；
- 成为支撑关键技术国际治理的全球技术标准生态系统的优胜者。建立行业和政府的能力，形成多方利益相关者联盟来制定技术标准；
- 利用政府的全景扫描能力，支持监管机构考虑新兴技术如何成为关键技术。行业和监管机构之间的对话将为英国的改革和促进技术发展提供信息。

## 3.3 金融领域的人工智能监管与影响

针对2023年3月发布的《促进创新的人工智能监管方法》白皮书，英国政府于2024年2月发布了回应。在此回应中，英格兰银行、审慎管理局、金融行为监管局以及其他已确定的监管机构被要求在4月底之前发布自己的人工智能监管计划。从其已发布的金融行业人工智能监管策略来看，总体而言，他们欢迎政府拟定的各部门引导，鼓励创新的方法。他们并未提出新的监管规则，英格兰银行、英国审慎管理局、英国金融行为监管局都认为他们已经设立了合适的监管框架来支撑政府拟定的原则。他们认为现有框架仍然适用于解决人工智能带来的风险，因为这些风险“并非唯一”。具体而言，监管机构计划依靠其商业原则、消费者责任、运营弹性规则手册、模型风险管理原则以及高级管理人员和认证制度 (SM&CR) 等工具。这些工具不仅会起到护栏作用，而且由于许多工具都以结果为中心，监管机构认为这将是相称的，并允许金融机构以安全的方式灵活适应和创新。

针对白皮书中提出的五项原则，监管机构概述了如何通过现有框架来贯彻这些原则：

- 安全性、保障性和稳健性：监管机构计划依靠其总体业务原则（例如，金融行为监管局的原则2，要求企业“以应有的技能、谨慎和勤勉的态度开展业务”）、基本条件以及高级管理系统和控制原始资料 (SYCS)。在更精细的层面上，运营弹性、第三方风险管理和外包框架也高度相关，企业有责任管理供应商可能影响重要商业服务的任何风险。
- 适当的透明度和可解释性：审慎管理局参考了其模型风险管理原则（其中明确要求考虑模型的可解释性和透明度）。金融行为监管局计划在消费者责任不适用的情况下，通过消费者责任（要求“与零售客户进行诚实和公开的交易”）和商业原则第7条（要求“适当考虑信息需求”）对此进行解释。英国GDPR规则也可能在公平透明处理个人数据方面发挥作用。
- 公平性：主要由金融行为监管局的消费者责任（例如，最终指导意见专门反对公司以放大偏见的方式使用人工智能）和其他关于公平对待弱势客户的指导意见来解决。金融行为监管局还参考了其商业原则，如原则6（适当考虑客户利益）、原则7（明确传达信息）、原则8（管理利益冲突）和原则9（建议的适用性），以及其消费者保护手册章节。所有监管机构还必须遵守《2010年平等法》，该法要求它们适当考虑对具有受保护特征的人消除歧视的必要性。

## 主要国家/地区人工智能监管

- 责任制与治理：该要求主要由高级管理人员和认证制度 (SM&CR) 涵盖。目前，人们一致认为，不需要定制SMCR AI角色，这一职责可以纳入现有职能中，例如首席运营职能 (SMF24) 和首席风险职能 (SMF4)。监管机构还参考了其规则手册中概述治理和组织要求的章节，而金融行为监管局也特别指出了消费者责任，审慎管理局也特别提出了模型风险管理的期望。
- 争议与纠偏：这主要通过金融行为监管局的投诉处理框架来解决。然而，也可以通过金融监察专员服务或金融服务补偿计划 (如适用) 提供更广泛的保护。
- 继续与其他监管机构和公司合作，包括通过数字监管合作论坛 (DRCF) 等，进一步监测和了解金融市场中的人工智能部署。例如，人工智能的发展也可能放大投机行为或更广泛的顺周期行为、增加网络风险以及增强互联性，从而对金融系统的稳定性构成威胁；
- 继续分析人工智能 (AI) /机器学习 (ML) 之外更广泛的技术发展的影响，例如量子计算的发展，以及这些发展如何相互作用；
- 继续参与国际倡议，如人工智能安全峰会、七国集团领导人关于广岛人工智能进程的声明、部长宣言、IOSCO、FSB、经合组织、GFIN和七国集团活动<sup>6</sup>。

从金融监管机构已公布的人工智能监管计划来看，他们近期将推进的举措有：

- 运行英国金融服务业机器学习调查的第三期；
- 考虑跟进人工智能公私论坛 (AIPPF) 成果，后续成立行业联盟；
- 考虑试点新型监管项目机会，例如人工智能沙盒；
- 建立在既定框架之上——例如，通过金融政策委员会 (FPC) 的网络压力测试和参与跨市场运营弹性小组 (CMORG) 来增强网络/运营弹性；
- 考虑各种选择，以解决当前数据治理和管理的分散监管局面；

然而，他们也认为，金融行业人工智能的发展与应用十分迅猛，需要密切关注。他们将继续监测金融机构的人工智能部署情况，以确定是否有必要对其框架进行任何修改。例如，如果这种方法限制而不是促进创新，或者不能充分保护消费者免受有意或无意的伤害，就需要重新考虑。

从对金融机构的影响来看，英国金融监管机构都利用了已有的监管体系与工具，为不同的人工智能供应商提供了足够的灵活性，以开发满足客户特定需求的创新应用。英国的监管方法是综合的、基于结果的，不同于其他司法管辖区（尤其是欧盟），后者建立了定制的、相对规范的规则手册。英国金融机构被要求在其现有的风险管理结构中考虑人工智能风险，而不是创造新的（潜在的孤立的）能力。因此，对金融机构而言，需要采取措施以确保已建立符合管理目标的人工智能风险管理工具，并充分吸纳已出台的要求<sup>7</sup>。

<sup>6-7</sup> 部分内容引用自毕马威国际文章《英国金融监管机构发布人工智能策略》，UK FS regulators publish their AI strategies - KPMG Global。



# 主要国家/地区人工智能监管

## 四、美国

### 4.1 美国体系概览

在美国，对人工智能的监管已经从自愿指导演变为行政行动（图12）。最初，2022年10月的《人工智能权利法案蓝图》奠定了基本的伦理考虑。随后，2023年1月，美国国家标准与技术研究院发布自愿性的《人工智能风险管理框架》。领先的人工智能公司于2023年7月做出了自愿承诺，表明行业准备应对人工智能的社会影响。2023年11月，《关于安全、可靠和值得信赖的人工智能的行政命令》标志着向监管监督的转变，该命令授权25个以上的机构解决与人工智能相关的危害，包括安全、隐私和歧视。这些机构的任务是制定规则、资助研究、评估风险，并通过人工智能开发人员的安全测试和报告来提高透明度。2024年5月，美国国家标准与技术研究院发布了两份生成式人工智能（GenAI）的指南草稿，以及降低合成内容风险的报告操作和人工智能标准全球合作的草稿，积极酝酿新的框架与指南。然而，在人工智能监管方面，联邦政府部门还没有采取重大的立法行动<sup>8</sup>。相比之下，州级别的立法活动较为活跃，自2020年开始，特别是2023年与2024年，部分州级人工智能立法已陆续落地生效。

从美国监管政策的重点来看，主要是以促进产业发展为主，在治理时侧重于维护公平和保护隐私。其治理模式是风险评估与风险管理框架。监管机制不是单一的依赖于行政监管，而是强调政府部门与私营部门的协作。并且，为了推动人工智能技术的发展，主张企业进行自我监管<sup>9</sup>。

图12：美国人工智能相关文件

## 2018

- 7月，美国财政部，《创造经济机会的金融体系：非银行金融、金融科技创新》

## 2019

- 2月，美国白宫，《关于保持美国在人工智能领域领导地位的行政命令》

## 2020

- 1月，美国伊利诺伊州《人工智能视频面试方案》生效
- 8月，美国保险专员协会，《人工智能原则》

## 2021

- 3月，美国财政部、货币监理署、联邦储备系统理事会、联邦存款保险公司、消费者金融保护局、全国信用合作社管理局，《金融机构人工智能应用（包含机器学习）的信息与反馈征集要求》
- 8月，美国证券交易委员会，《数字化参与实践（DEPs）信息征集要求》

## 2022

- 1月，美国联邦贸易委员会，保障规则后修正案生效
- 5月，美国消费者金融保护局，《关于复杂算法不当行动要求的通知》
- 8月，美国联邦贸易委员会，《商业监控与数据安全的拟议规则预先通知》
- 10月，美国白宫，《人工智能权利法案蓝图》

## 2023

- 1月，美国国家标准与技术研究院，《人工智能风险管理自愿性框架》
- 1月，加利福尼亚州，CPRA/CPRA 正式生效
- 1月，弗吉尼亚州，《消费者数据保护法》（NCDDPA）生效
- 3月，美国证券交易委员会，修改法规S-P的提案
- 4月，美国商务部国家电信和信息管理局，《人工智能的保证与信任》征求意见稿
- 4月，美国消费者金融保护局、司法部民权司、平等就业机会委员会、联邦贸易委员会，《关于监管和执行“自动化系统”的机构间声明》
- 5月，美国联邦贸易委员会，《关于通用生物识别信息的政策声明》
- 5月，美国国会，多场关于人工智能、监管失败、知识产权的听证会
- 5月，犹他州，《消费者数据保护法》生效
- 6月，美国货币监理署、联邦储备系统理事会、联邦存款保险公司、全国信用合作社管理局、消费者金融保护局、联邦住房金融局，《关于自动化估值模型的金融机构改革救济法修正案通知》
- 7月，美国证券交易委员会，《关于预测数据分析利益冲突的拟议通知》
- 7月，纽约州，自动化就业决策工具法生效
- 7月，科罗拉多州，科罗拉多州隐私法生效
- 7月，康涅狄格州，康涅狄格州数据隐私法生效
- 7月，康涅狄格州，关于人工智能、自动化决策和个人数据隐私的法案
- 9月，美国消费者金融保护局，《贷款人使用人工智能拒绝信贷的指南》
- 9月，美国网络安全和基础设施安全局，《深度伪造技术的网络安全信息表》
- 9月，美国白宫，《人工智能公司关于人工智能风险管理的自愿承诺情况说明书》
- 10月，美国白宫，《关于安全、可靠和值得信赖的人工智能的行政命令》
- 12月，美国国家标准与技术研究院，《协助履行关于安全、可靠和值得信赖的人工智能开发和使用的行政命令下职责的信息请求》

## 2024

- 1月，美国白宫，成立关于人工智能的金融委员会两党工作组
- 1月，美国商品期货交易委员会，《人工智能在商品期货市场应用的意见征询》
- 1月，美国白宫，《关于行政命令人工智能相关举措的情况说明书》
- 1月，纽约州金融服务部，《关于保险业人工智能的建议指南》
- 3月，犹他州，《人工智能政策法案》
- 5月，美国两院参议院人工智能工作组，《推动美国人工智能创新》
- 5月，美国国家标准与技术研究院，《指南：人工智能风险管理框架：生成式人工智能（GenAI）简介（NIST AI 600-1）》
- 5月，美国国家标准与技术研究院，《指南：生成式人工智能（GenAI）和两用基础模型的安全软件开发实践（NIST特别出版物（SP）800-218A）》
- 5月，美国国家标准与技术研究院，《报告：降低合成内容带来的风险（NIST AI 100-4）》
- 5月，美国国家标准与技术研究院，《计划：全球人工智能标准合作（NIST AI 100-5）》
- 6月，科罗拉多州，科罗拉多州人工智能法案（CAIA）
- 6月，美国金融业监管局，《监管通知24-09：FINRA提醒成员在使用生成式人工智能和大型语言模型时的监管义务》
- 6月，美国财政部，《关于人工智能在金融服务领域的应用、机遇和风险信息请求》

说明：州立法/文件 联邦政府部门及其他组织立法/文件 纳入重点分析文件

信息来源：毕马威自行收集整理。

<sup>8</sup> 部分内容引用自Bank for International Settlements (BIS) Working Papers No. 1194 Intelligent financial system: how AI is transforming finance.

<sup>9</sup> 引用自《人工智能监管制度概览》，国家社会科学基金重点项目“人工智能研发与应用风险治理的财税法协同机制研究”（立项编号：21AFX021）的阶段性成果，来源人民法院报。

# 主要国家/地区人工智能监管

## 4.2 重要监管文件分析

以下选择美国具有标志性意义的三份监管文件进行分析。

### 《人工智能权利法案蓝图》（2022年10月）

2022年10月4日，白宫科技政策办公室（OSTP）发布《人工智能权利法案蓝图》（以下简称“法案蓝图”）。《法案蓝图》由五项原则和相关实践组成，旨在指导自动化系统的设计、使用和部署，以保护人工智能时代美国公众的权利。这些原则是通过与美国公众广泛协商制定的，是构建和部署与民主价值观相一致并保护公民权利、公民自由和隐私的自动化系统的蓝图。从具体内容来看，《法案蓝图》包括前言、五项原则（图13）、关于应用《法案蓝图》的说明，以及一份技术指南，其中给出了从各级政府到各种规模的公司等多种组织可以采取的具体步骤，以维护这些价值观。尽管来自私营部门、政府和国际财团的专家已经发布了指导负责任地使用自动化系统的原则和框架；《法案蓝图》的框架提供了一个与行业无关的国家价值观声明和工具包，为将这些保护纳入政策、实践或技术设计过程提供信息<sup>10</sup>。

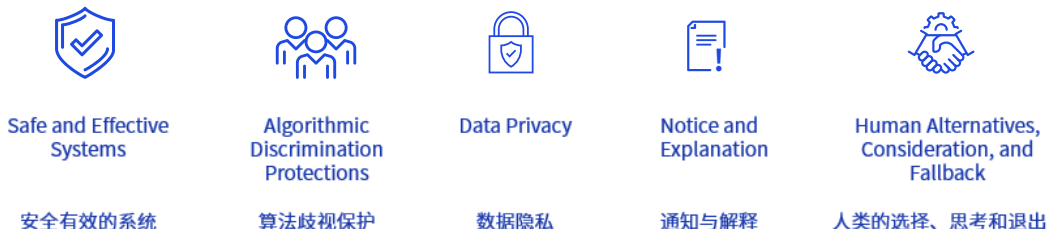
在现有法律或政策（如特定行业的隐私法和监督要求）尚未提供指导的情况下，应使用《法案蓝图》来为政策决策提供信息。

### 《人工智能风险管理框架》（2023年1月）

2023年1月26日，美国国家标准与技术研究院（隶属于美国商务部）发布了《人工智能风险管理框架（AI RMF 1.0）》（以下简称“风险框架”），旨在帮助设计、开发、部署或使用人工智能系统的组织提高人工智能风险管理的能力，并促进人工智能系统的可信度和负责任的开发和使用的。《风险框架》是一个自愿性的技术框架，不具备强制效力。

《风险框架》分为两部分，第一部分讨论了组织如何构建与人工智能相关风险的框架，并描述了目标受众（图14）。之后分析了人工智能相关风险和可信度，概述了值得信赖的人工智能系统的特征，包括有效和可靠（Valid and Reliable），安全（Safe），保障和有弹性（Secure and Resilient），负责和透明（Accountable and Transparent），可解释和可交互解释（Explainable & Interpretable），隐私增强（Privacy-Enhanced），以及公平——对有害偏见进行

图13：《人工智能权利法案蓝图》的5项原则



信息来源：毕马威根据白宫科技政策办公室官网信息翻译整理。

<sup>10</sup> 部分内容引用自美国白宫科技政策办公室，Blueprint for an AI Bill of Rights的框架简介。

# 主要国家/地区人工智能监管

图14：人工智能系统的全生命周期与关键维度



信息来源：毕马威根据《人工智能风险管理框架》翻译整理。

管理 (Fair-With Harmful Bias Managed) 。

第二部分包括《风险框架》的“核心”。它描述了四项具体职能 (图15)，以帮助组织在实践中应对人工智能系统的风险，即治理

( Govern ) 、映射 ( Map ) 、测量 ( Measure ) 和管理 ( Manage ) ，并进一步细分了四个职能的类别和子类别。其中，治理适用于组织人工智能风险管理流程和程序的所有阶段，映射、测量与管理功能可以应用于人工智能系统特定的环境和人工智能生命周期的特定阶段。

# 主要国家/地区人工智能监管

图15：治理、映射、测量、管理四项核心职能



信息来源：毕马威根据《人工智能风险管理框架》翻译整理。

## 《关于安全、可靠和值得信赖的人工智能的行政命令》（2023年10月）

2023年10月30日，美国白宫发布拜登签署的最新行政命令《关于安全、可靠和值得信赖的人工智能的行政命令》（以下简称“行政命令”），呼吁采取“整个政府”的方法来建立安全、可靠和值得信赖的人工智能，强调人工智能开发和部署的透明度和问责制，并指示联邦机构优先考虑促进经济增长和创造就业机会的人工智能技术，同时尽量减少对工人和社区的潜在负面影响。《行政命令》遵循政府的《人工智能权利法案蓝图》，是一套指导设计、使用和部署自动化系统以保护公众的原则和做法<sup>11</sup>。《行政命令》提出了8项指导原则和优先事项（图16），同时也对

美政府下一步的行動措施与职责分工进行了部署与安排。

## 4.3 金融领域的人工智能监管与影响

截止目前，美国金融领域的人工智能监管规则较少，具有强制性遵循意义的两份要求来自于美国消费者金融保护局（CFPB）发布的《关于复杂算法不当行动要求的通知》与《贷款人使用人工智能拒绝信贷的指南》，旨在聚焦具体应用场景，保护金融消费者公平，避免算法与人工智能技术带来的系统性歧视。美国财政部（US Treasury）、美国证券交易委员会（SEC）、美国联邦贸易委员会（FTC）、美国商务部（DOC）、美国商品期货交易委员会（CFTC）等金融监管机构发布

<sup>11</sup> 部分内容引用自毕马威美国发布的监管要求报告，AI Regulations: Present & Future, Apr. 2024。



## 主要国家/地区人工智能监管

的文件主要形式为征求意见 (Request for Comment, RFC)、拟议规则预先通知 (Advance Notice of Proposed Rulemaking, ANPR)、拟议规则通知 (Notice of Proposed Rulemaking, NPR)、信息请求 (Request for Information, RFI)、

研究报告等, 主题涉及商业监控与数据安全, 人工智能运用于鉴证与信托, 自动化估值模型, 使用预测数据分析的利益冲突等, 相关内容尚处于公众探讨与意见征询阶段, 对金融机构不具备约束力, 但预期将在未来陆续落地出台。



# 主要国家/地区人工智能监管

图16: 《关于安全、可靠和值得信赖的人工智能的行政命令》的8项指导原则与优先事项

01

人工智能必须是安全的和有保障的。  
Artificial Intelligence must be safe and secure.

02

促进负责任的创新、竞争和合作将使美国在人工智能领域处于领先地位，并释放该技术的潜力，解决社会上一些最困难的挑战。  
Promoting responsible innovation, competition, and collaboration will allow the United States to lead in AI and unlock the technology's potential to solve some of society's most difficult challenges.

03

负责任地开发和使用人工智能需要承诺支持美国工人。  
The responsible development and use of AI require a commitment to supporting American workers.

04

人工智能政策必须与我的政府致力于促进公平和公民权利的承诺保持一致。  
Artificial intelligence policies must be consistent with my Administration's dedication to advancing equity and civil rights.

05

必须保护在日常生活中越来越多地使用、互动或购买人工智能和人工智能产品的美国人的利益。  
The interests of Americans who increasingly use, interact with, or purchase AI and AI-enabled products in their daily lives must be protected.

06

随着人工智能的不断进步，美国人的隐私和公民自由必须得到保护。  
Americans' privacy and civil liberties must be protected as AI continues advancing.

07

重要的是要管理联邦政府自身使用人工智能的风险，并提高其监管、治理和支持负责任地使用人工智能以为美国人提供更好结果的内部能力。  
It is important to manage the risks from the Federal Government's own use of AI and increase its internal capacity to regulate, govern, and support responsible use of AI to deliver better results for Americans.

08

联邦政府应该引领全球社会、经济和技术进步，就像美国在以前的颠覆性创新和变革时代一样。  
The Federal Government should lead the way to global societal, economic, and technological progress, as the United States has in previous eras of disruptive innovation and change.

数据来源：毕马威根据《关于安全、可靠和值得信赖的人工智能的行政命令》翻译整理。

# 主要国家/地区人工智能监管

## 五、中国

### 5.1 中国体系概览

中国较早就启动了人工智能相关的规划与治理活动（图17）。2017年7月国务院发布《新一代人工智能发展规划》（以下简称“规划”），并在第五部分“保障措施”中，作出了制定促进人工智能发展的法律法规和伦理规范，完善支持人工智能发展的重点政策，建立人工智能技术标准和知识产权体系，建立人工智能安全监管和评估体系等整体部署，并明确由国家科技体制改革和创新体系建设领导小组牵头统筹协调，推动人工智能相关法律法规建设。承接规划部署，为促进新一代人工智能健康发展，加强人工智能法律、伦理、社会问题研究，积极推动人工智能全球治理，新一代人工智能发展规划推进办公室成立了国家新一代人工智能治理专业

委员会，并陆续于2019年6月和2021年9月推出了《新一代人工智能治理原则——发展负责任的人工智能》与《新一代人工智能伦理规范》，明确了中国人工智能治理原则与伦理规范。中国亦积极参与国际人工智能治理，于2023年发布了《全球人工智能治理倡议》，并在2024年7月召开的2024世界人工智能大会暨人工智能全球治理高级别会议上，发布了《人工智能全球治理上海宣言》。

整体来看，目前中国的人工智能领域具有强制约束力的监管文件较少，已出台文件主要集中于规划蓝图、治理原则、伦理规范等软性约束。但预计未来将有更多人工智能领域的法律法规和政策文件陆续出台。

图17：中国人工智能相关文件



信息来源：毕马威自行收集整理。

# 主要国家/地区人工智能监管

## 5.2 重要监管文件分析

以下选择中国具有标志性意义的三份监管文件进行分析。

### 《新一代人工智能治理原则——发展负责任的人工智能》（2019年6月）

2019年6月17日，国家新一代人工智能治理专业委员会发布《新一代人工智能治理原则——发展负责任的人工智能》（以下简称“治理原则”），提出了人工智能治理的框架和行动指南。这是中国促进新一代人工智能健康发展，加强人工智能法律、伦理、社会问题研究，积极推动人工智能全球治理的一项重要成果<sup>12</sup>。《治理原则》包含8个方面（图18），具体内容如下：

- 和谐友好。人工智能发展应以增进人类共同福祉为目标；应符合人类的价值观和伦理道德，促进人机和谐，服务人类文明进步；应以保障社会安全、尊重人类权益为前提，避免误用，禁止滥用、恶用；

- 公平公正。人工智能发展应促进公平公正，保障利益相关者的权益，促进机会均等。通过持续提高技术水平、改善管理方式，在数据获取、算法设计、技术开发、产品研发和应用过程中消除偏见和歧视；
- 包容共享。人工智能应促进绿色发展，符合环境友好、资源节约的要求；应促进协调发展，推动各行各业转型升级，缩小区域差距；应促进包容发展，加强人工智能教育及科普，提升弱势群体适应性，努力消除数字鸿沟；应促进共享发展，避免数据与平台垄断，鼓励开放有序竞争；
- 尊重隐私。人工智能发展应尊重和保护个人隐私，充分保障个人的知情权和选择权。在个人信息的收集、存储、处理、使用等各环节应设置边界，建立规范。完善个人数据授权撤销机制，反对任何窃取、篡改、泄露和其他非法收集利用个人信息的行为；

图18：新一代人工智能治理8项原则



信息来源：毕马威根据《新一代人工智能治理原则——发展负责任的人工智能》整理。

12、部分内容引用自人民日报，我国发布《治理原则》发展负责任的人工智能，2019年6月18日。



## 主要国家/地区人工智能监管

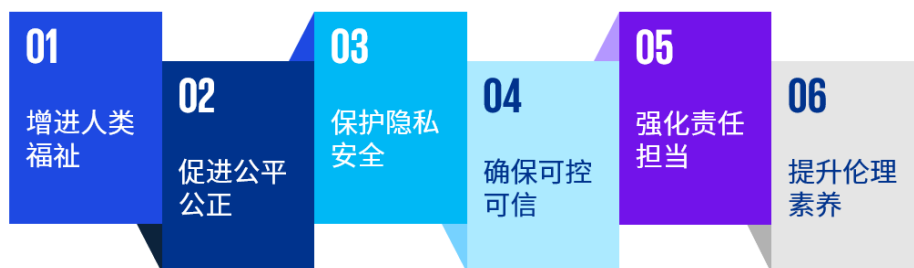
- 安全可控。人工智能系统应不断提升透明性、可解释性、可靠性、可控性，逐步实现可审核、可监督、可追溯、可信赖。高度关注人工智能系统的安全，提高人工智能鲁棒性及抗干扰性，形成人工智能安全评估和管控能力；
- 共担责任。人工智能研发者、使用者及其他相关方应具有高度的社会责任感和自律意识，严格遵守法律法规、伦理道德和标准规范。建立人工智能问责机制，明确研发者、使用者和受用者等的责任。人工智能应用过程中应确保人类知情权，告知可能产生的风险和影响。防范利用人工智能进行非法活动；
- 开放协作。鼓励跨学科、跨领域、跨地区、跨国界的交流合作，推动国际组织、政府部门、科研机构、教育机构、企业、社会组织、公众在人工智能发展与治理中的协调互动。开展国际对话与合作，在充分尊重各国人工智能治理原则和实践的前提下，推动形成具有广泛共识的国际人工智能治理框架和标准规范；
- 敏捷治理。尊重人工智能发展规律，在推动人工智能创新发展、有序发展的同时，及时发现和解决可能引发的风险。不断提升智能化技术手段，优化管理机制，完善治理体系，推动治理原则贯穿人工智能产品和服务的全生命周期。对未来更高级人工智能的潜在风险持续开展研究和预判，确保人工智能始终朝着有利于社会的方向发展。

### 《新一代人工智能伦理规范》（2021年9月）

2021年9月25日，国家新一代人工智能治理专业委员会发布了《新一代人工智能伦理规范》(以下简称《伦理规范》)，旨在将伦理道德融入人工智能全生命周期，为从事人工智能相关活动的自然人、法人和其他相关机构等提供伦理指引<sup>13</sup>。《伦理规范》提出了6项基本伦理要求，以及人工智能管理、研发、供应、使用等特定活动的18项具体伦理要求。

6项基本伦理要求（图19）为：

图19：新一代人工智能基本伦理要求



信息来源：毕马威根据《新一代人工智能伦理规范》整理。

13、引用自科学技术部官方网站，《新一代人工智能伦理规范》发布，2021年9月26日。

## 主要国家/地区人工智能监管

- 增进人类福祉。坚持以人为本，遵循人类共同价值观，尊重人权和人类根本利益诉求，遵守国家或地区伦理道德。坚持公共利益优先，促进人机和谐友好，改善民生，增强获得感幸福感，推动经济、社会及生态可持续发展，共建人类命运共同体；
  - 促进公平公正。坚持普惠性和包容性，切实保护各相关主体合法权益，推动全社会公平共享人工智能带来的益处，促进社会公平正义和机会均等。在提供人工智能产品和服务时，应充分尊重和帮助弱势群体、特殊群体，并根据需要提供相应替代方案；
  - 保护隐私安全。充分尊重个人信息知情、同意等权利，依照合法、正当、必要和诚信原则处理个人信息，保障个人隐私与数据安全，不得损害个人合法数据权益，不得以窃取、篡改、泄露等方式非法收集利用个人信息，不得侵害个人隐私权；
  - 确保可控可信。保障人类拥有充分自主决策权，有权选择是否接受人工智能提供的服务，有权随时退出与人工智能的交互，有权随时中止人工智能系统的运行，确保人工智能始终处于人类控制之下；
  - 强化责任担当。坚持人类是最终责任主体，明确利益相关者的责任，全面增强责任意识，在人工智能全生命周期各环节自省自律，建立人工智能问责机制，不回避责任审查，不逃避应负责任；
  - 提升伦理素养。积极学习和普及人工智能伦理知识，客观认识伦理问题，不低估不夸大伦理风险。主动开展或参与人工智能伦理问题讨论，深入推动人工智能伦理治理实践，提升应对能力。
- 18项人工智能管理、研发、供应、使用等特定活动的具体伦理要求为（图20）：

图20：新一代人工智能管理、研发、供应、使用的伦理要求



信息来源：毕马威根据《新一代人工智能伦理规范》整理。

# 主要国家/地区人工智能监管

## 《全球人工智能治理倡议》（2023年10月）

2023年10月18日，中国国家互联网信息办公室发布《全球人工智能治理倡议》（以下简称“倡议”），围绕人工智能发展、安全、治理三方面系统阐述了人工智能治理中国方案。《倡议》提出，各国应秉持共同、综合、合作、可持续的安全观，坚持发展和安全并重的原则，通过对话与合作凝聚共识，构建开放、公正、有效的治理机制，促进人工智能技术造福于人类，推动构建人类命运共同体。各国应在人工智能治理中加强信息交流和技术合作，共同做好风险防范，形成具有广泛共识的人工智能治理框架和标准规范，不断提升人工智能技术的安全性、可靠性、可控性、公平性。并进而提出了11项倡议：

- 发展人工智能应坚持“以人为本”理念，以增进人类共同福祉为目标，以保障社会安全、尊重人类权益为前提，确保人工智能始终朝着有利于人类文明进步的方向发展。积极支持以人工智能助力可持续发展，应对气候变化、生物多样性保护等全球性挑战；
- 面向他国提供人工智能产品和服务时，应尊重他国主权，严格遵守他国法律，接受他国法律管辖。反对利用人工智能技术优势操纵舆论、传播虚假信息，干涉他国内政、社会制度及社会秩序，危害他国主权；
- 发展人工智能应坚持“智能向善”的宗旨，遵守适用的国际法，符合和平、发展、公平、正义、民主、自由的全人类共同价值，共同防范和打击恐怖主义、极端势力和跨国组织犯罪集团对人工智能技术的恶意滥用。各国尤其是大国对在军事领域研发和使用人工智能技术应该采取慎重负责的态度；
- 发展人工智能应坚持相互尊重、平等互利的原则，各国无论大小、强弱，无论社会制度如何，都有平等发展和利用人工智能的权利。鼓励全球共同推动人工智能健康发展，共享人工智能知识成果，开源人工智能技术。反对以意识形态划线或构建排他性集团，恶意阻挠他国人工智能发展。反对利用技术垄断和单边强制措施制造发展壁垒，恶意阻断全球人工智能供应链；
- 推动建立风险等级测试评估体系，实施敏捷治理，分类分级管理，快速有效响应。研发主体不断提高人工智能可解释性和可预测性，提升数据真实性和准确性，确保人工智能始终处于人类控制之下，打造可审核、可监督、可追溯、可信赖的人工智能技术；
- 逐步建立健全法律和规章制度，保障人工智能研发和应用中的个人隐私与数据安全，反对窃取、篡改、泄露和其他非法收集利用个人信息的行为；
- 坚持公平性和非歧视性原则，避免在数据获取、算法设计、技术开发、产品研发与应用过程中，产生针对不同或特定民族、信仰、国别、性别等偏见和歧视；
- 坚持伦理先行，建立并完善人工智能伦理准则、规范及问责机制，形成人工智能伦理指南，建立科技伦理审查和监管制度，明确人工智能相关主体的责任和权力边界，充分尊重并保障各群体合法权益，及时回应国内和国际相关伦理关切；

## 主要国家/地区人工智能监管

- 坚持广泛参与、协商一致、循序渐进的原则，密切跟踪技术发展形势，开展风险评估和政策沟通，分享最佳操作实践。在此基础上，通过对话与合作，在充分尊重各国政策和实践差异性基础上，推动多利益攸关方积极参与，在国际人工智能治理领域形成广泛共识；
- 积极发展用于人工智能治理的相关技术开发与应用，支持以人工智能技术防范人工智能风险，提高人工智能治理的技术能力；
- 增强发展中国家在人工智能全球治理中的代表性和发言权，确保各国人工智能发展与治理的权利平等、机会平等、规则平等，开展面向发展中国家的国际合作与援助，不断弥合智能鸿沟和治理能力差距。积极支持在联合国框架下讨论成立国际人工智能治理机构，协调国际人工智能发展、安全与治理重大问题。

### 5.3 金融领域的人工智能监管与影响

目前中国尚未出台金融领域人工智能监管的专门法律法规，但在消费者保护领域，已有算法规制、互联网广告等相关联的适用监管要求。在行业标准与自律规范方面，中国人民银行于2021年3月发布了《人工智能算法金融应用评价规范》，提供了金融领域应用人工智能算法的基本要求、评价方法和判定准则。2022年10月，中国银行业协会发布《人工智能模型风险管理框架》，规范了银行业金融机构在开展人工智能模型需求分析、数据准备、模型构建、检验验证、模型部署、模型监控、持续验证与确认、模型修正、模型下线等关键活动过程中进行风险管理需考虑的相关方面。2023年8月，国家市场监督管理总局、国家标准化委员会发布《信息

安全技术机器学习算法安全评估规范》(GB/T 42888-2023)，规定了机器学习算法技术和服务的\*\*安全要求和评估方法，以及机器学习算法安全评估流程等。2023年9月，中国信息通信研究院联合多家企业发布了国内首个金融行业大模型标准《面向行业的大规模预训练模型技术和应用评估方法 第1部分：金融大模型》，旨在规范金融大模型的研发和应用过程，提高模型的准确性和可靠性，促进金融行业的智能化发展。但相关规范均为参考、指南类的软性文件，对金融机构不具有强制约束力。

以下我们将对《人工智能算法金融应用评价规范》、《人工智能模型风险管理框架》两份金融业人工智能相关文件进行重点分析。

#### 《人工智能算法金融应用评价规范》(2021年3月)

为全面提升人工智能技术在金融领域应用和管理水平，推动金融与科技深度融合协调发展，2021年3月26日，中国人民银行发布金融行业标准 JR/T 0221-2021《人工智能算法金融应用评价规范》(以下简称“规范”)。《规范》针对人工智能技术应用存在的算法黑箱、算法同质化、模型缺陷等潜在风险问题，建立了人工智能金融应用算法评价框架，从安全性、可解释性、精准性和性能等方面系统化地提出基本要求、评价方法和判定准则，为金融机构加强智能算法应用风险管理提供了指引<sup>14</sup>。《规范》明确其适用对象为开展人工智能算法金融应用的金融机构、算法提供商和第三方安全评估机构等。

首先，《规范》提出了整体评价框架(图21)，并区分各项内容的适用场景为资金类场景和非资金类场景<sup>15</sup>。

<sup>14</sup> 部分内容引用自经济参考网文章《<人工智能算法金融应用评价规范>出炉 助力打造金融创新发展新引擎》。

<sup>15</sup> 引用内容来源于中国人民银行，《人工智能算法金融应用评价规范》(JR/T 0221-2021)。



# 主要国家/地区人工智能监管

图21: AI算法评价内容框架图



信息来源: 引用自《人工智能算法金融应用评价规范》。

其次,《规范》对可采用的评价方法进行了详细列示与说明,包括查阅资料、查看系统、访谈人员、系统测试、攻击测试、算法测试与查看算法。

最后,《规范》结合整体框架提出的评价内容,对每项内容包含的基本要求,建议采用的评价方法及判定准则进行了详细阐述。并以附录形式提供了金融行业AI精准性的相关指标定义,覆盖分类算法模型、回归算法模型、图像识别模型、自然语言处理模型等四类模型的15个指标。

## 《人工智能模型风险管理框架》(2022年10月)

2022年10月24日,中国银行业协会发布团体技术报告 TR/CBA 106-2022《人工智能模型风险管理框架》(以下简称“框架”)。

《框架》针对银行业金融机构在管理、研发、供应、使用人工智能模型过程中面临的主要风险挑战,根据行业特性,给出银行业金融机构在开展人工智能模型需求分析、数据准备、模型构建、检验验证、模型部署、模型监控、持续验证与确认、模型修正、模型下线等关键活动过程中进行风险管理需要考虑内容的框架。《框架》适用于银行业金融机构,向银行业金融机构提供人工智能服务和支持的相关机构亦能作为参考<sup>16</sup>。

<sup>16</sup> 部分内容引用自团体技术报告 TR/CBA 106-2022《人工智能模型风险管理框架》。

# 主要国家/地区人工智能监管

《框架》采取了“分级管理、逐步递增”的基本策略，将人工智能模型风险管理措施分为逐步递增的三个级别（图22）。根据其内容来看，第一级为“不能有效地管控风险”，第二级为人工智能风险管理的基础要求，第三级为较佳实践。

《框架》依据模型生命周期中的九个主要环节，以及特殊过程的两类事项，详细阐述了其适用的风险分级管理措施（图23），内容详细覆盖组织架构与职责分工、流程与工作机制、报告文档规范、系统平台与工具功能、模型监控指标等，对于银行业金融机构具有较强的指导意义。

图22：人工智能风险管理措施的三个级别



信息来源：毕马威根据《人工智能模型风险管理框架》整理。

图23：模型生命周期与特殊过程



信息来源：毕马威根据《人工智能模型风险管理框架》整理。

# 主要国家/地区人工智能监管

从对金融机构的影响与合规建议来看，整体而言，金融机构面临的人工智能相关硬性约束较少，而已发布的软性指导文件集中在各行业通用性的、高层次的人工智能治理原则与伦理规范；针对金融业侧重于算法模型开发应用过程中的风险应对、安全评估等技术领域的规范要求；以及消费者权益保护视角的算法规制与互联网广告要求等。

但随着中国人工智能技术的快速发展及应用转化，金融行业已广泛将人工智能技术投入于智能营销、智慧投顾、投诉处理、风险管

理、制度审阅、合规测试等领域的探索。建议金融机构一方面做好已发布的监管规则的合规，以及结合机构管理实际，主动应用各项自律规范与行业标准；另一方面主动调研行业市场及机构内部的人工智能应用现状，保持对于人工智能技术与监管立法进程的动态关注，并结合风险控制目标，前瞻性地将人工智能治理要求与风险管控融入日常管理流程中（图24）。同时关注机构内部人工智能相关人才与资源储备，以备未来应对人工智能的应用发展与监管变化。

图24：中国人工智能相关要求与金融机构合规建议

分类	立法机构	文件名称	金融机构合规建议
行业通用 (非强制)	国务院	《新一代人工智能发展规划》	参考搭建本机构适用的人工智能治理框架、管理流程等
	国家新一代人工智能治理专业委员会	《新一代人工智能治理原则——发展负责任的人工智能》	
	国家新一代人工智能治理专业委员会	《新一代人工智能伦理规范》	
	国家互联网信息办公室	Global AI Governance Initiative全球人工智能治理倡议	
	2024世界人工智能大会暨人工智能全球治理高级别会议	《人工智能全球治理上海宣言》	
行业通用 (强制)	全国信息安全标准化技术委员会	关于发布《网络安全标准实践指南——生成式人工智能服务内容标识方法》的通知	落实合规遵循，做好外规内化
	国家互联网信息办公室，国家发展和改革委员会，教育部，科学技术部，工业和信息化部，公安部，国家广播电视总局	《生成式人工智能服务管理暂行办法》	
	国家市场监督管理总局	互联网广告管理办法	
金融业适用 (非强制)	国家市场监督管理总局	互联网信息服务算法推荐管理规定	在模型全生命周期管理、算法应用、金融大模型训练等领域，根据本机构实际情况，主动应用行业规范与领先实践标准
	中国人民银行	《人工智能算法金融应用评价规范》	
	中国银行业协会	《人工智能模型风险管理框架》	
技术标准	中国信息通信研究院	《面向行业的大规模预训练模型技术和应用评估方法 第1部分：金融大模型》	在相关技术领域应用技术标准
	国家市场监督管理总局、国家标准化管理委员会	《信息安全技术机器学习算法安全评估规范 (GB/T 42888-2023)》	
	国家市场监督管理总局、国家标准化管理委员会	人工智能 情感计算用户界面 模型 (GB/T 40691-2021)	
	国家市场监督管理总局、国家标准化管理委员会	信息技术 人工智能 术语 (标准号: GB/T 41867-2022)	
	国家市场监督管理总局、国家标准化管理委员会	人工智能 知识图谱技术框架 (GB/T 42131-2022)	
国家市场监督管理总局、国家标准化管理委员会	信息技术 人工智能 平台计算资源规范 (GB/T 42018-2022)		

数据来源：毕马威自行收集整理。

## 结语

总体而言，在人工智能监管方面，国际组织与各国家地区截至目前出台的专门性法律法规较少，其立法尚处于研究摸索与实践探寻的初级阶段；但随着欧盟《人工智能法案》的落地生效，预计未来各国家/地区的监管立法活动将提速。目前多数已发布文件为人工智能相关的监管指导性文件，其治理框架、原则、伦理规范等，对金融机构具有一定的指导意义，但政府部门与监管机构仍将大部分实际实施空间留给了金融机构自己。

随着人工智能技术在金融业的应用不断拓展与丰富，其蕴含的风险与问题将不断涌现暴露，特别是在金融业整体严监管态势与结果合规的导向下，需要金融机构主动作为，积

极应对，一方面贯彻已颁布的监管要求与规范倡导，另一方面运用机构已有的风险三道防线设置与管理机制，以及与人工智能相关的数据安全、个人信息保护、网络安全等领域管理机制与管理工具，从风险视角实施目标管控，确保在整个端到端商业模式中满足监管预期并充分考虑人工智能风险。

此外，考虑到国际不同区域、国家之间人工智能监管立法发展不均衡的现状，全球化经营的大型金融机构在整体开发与部署人工智能时，还需关注不同区域/国家机构的监管要求与合规达标，特别是欧盟等人工智能立法步速较快的国家与地区，以充分挖掘人工智能技术的金融应用潜力，在合理管控潜在风险的前提下促进创新。



## 附件一：全球人工智能监管文件清单

序号	立法机构	文件名称	年度	具体时间	国别/区域
1	国务院	《新一代人工智能发展规划》	2017	Jul.2017	中国
2	25 European Countries	Declaration of cooperation on Artificial Intelligence (AI)	2018	Apr.2018	欧洲
3	Federal Financial Supervisory Authority of Germany (BaFin)	Big Data meets AI	2018	Jul.2018	德国
4	US Treasury	A financial system that creates economic opportunities: nonbank financials, fintech, and innovation	2018	Jul.2018	美国
5	Monetary Authority of Singapore (MAS)	Principles to promote fairness, ethics, accountability and transparency (FEAT) in the use of AI and data analytics in Singapore's financial sector	2018	Nov.2018	新加坡
6	Financial Sector Supervisory Commission of Luxembourg (CSSF)	AI: Opportunities, risks and recommendations for the financial sector	2018	Dec.2018	卢森堡
7	the European Commission	Coordinated Plan on Artificial Intelligence	2018	Dec.2018	欧盟
8	White House	Executive Order on Maintaining American Leadership in AI	2019	Feb.2019	美国
9	Independent High-level Expert Group on AI (set up by the European Commission)	Ethics guidelines for trustworthy AI (draft version in Dec. 2018)	2019	Apr.2019	欧盟
10	European Insurance and Occupational Pensions Authority	Big data analytics in motor and health insurance: A thematic review	2019	May.2019	欧盟
11	Organisation for Economic Cooperation and Development (OECD)	AI Principles	2019	May.2019	国际组织
12	G20	AI Principles	2019	Jun.2019	国际组织
13	国家新一代人工智能治理专业委员会	《新一代人工智能治理原则——发展负责任的人工智能》	2019	Jun.2019	中国
14	Netherlands Bank	General principles for the use of AI in the financial sector	2019	Jul.2019	荷兰
15	Bank of England and Financial Conduct Authority	Research Note: Machine Learning in UK financial services	2019	Oct.2019	英国
16	Hong Kong Monetary Authority	High-level principles on AI	2019	Nov.2019	香港
17	Hong Kong Monetary Authority	Consumer protection in respect of Use of Big Data Analytics and Artificial Intelligence by Authorized Institutions	2019	Nov.2019	香港
18	Illinois	AI Video Interview Act became effective	2020	Jan.2020	美国
19	European Banking Authority	Report on big data and advanced analytics	2020	Jan.2020	欧盟

序号	立法机构	文件名称	年度	具体时间	国别/区域
20	French Prudential Supervision and Resolution Authority (ACPR)	Governance of AI in Finance	2020	Jan.2020	法国
21	UK's Information Commissioner's Office	draft Guidance on the AI auditing framework	2020	Feb.2020	英国
22	the European Commission	White paper on AI: a European approach to excellence and trust	2020	Feb.2020	欧盟
23	UK's Information Commissioner's Office	Guidance on AI and data protection	2020	Jul.2020	英国
24	National Association of Insurance Commissioners	Principles for Artificial Intelligence	2020	Aug.2020	美国
25	Department of the Treasury, Office of the Comptroller of the Currency, Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, Bureau of Consumer Financial Protection, National Credit Union Administration	Request for information and comment on financial institutions' use of AI, including machine learning	2021	Mar.2021	美国
26	中国人民银行	《人工智能算法金融应用评价规范》	2021	Mar.2021	中国
27	the European Commission	Proposal for a regulation laying down harmonised rules on AI	2021	Apr.2021	欧盟
28	European Insurance and Occupational Pensions Authority	Artificial intelligence governance principles: towards ethical and trustworthy artificial intelligence in the European insurance sector	2021	Jun.2021	欧盟
29	Federal Financial Supervisory Authority of Germany (BaFin)	Big data and artificial intelligence: Principles for the use of algorithms in decision-making processes	2021	Jun.2021	德国
30	Securities and Exchange Commission (SEC)	Request for information on Digital engagement practices (DEPs)	2021	Aug.2021	美国
31	Bank for International Settlements (BIS)	Humans keeping AI in check – emerging regulatory expectations in the financial sector	2021	Aug.2021	国际组织
32	International Organization of Securities Commissions (IOSCO)	The use of artificial intelligence and machine learning by market intermediaries and asset managers Final Report	2021	Sep.2021	国际组织
33	国家新一代人工智能治理专业委员会	《新一代人工智能伦理规范》	2021	Sep.2021	中国
34	Department for Science, Innovation and Technology, Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy	National AI Strategy	2021	Sep.2021	英国

序号	立法机构	文件名称	年度	具体时间	国别/区域
35	Monetary Authority of Singapore (MAS)	National programme to deepen AI capabilities in financial services	2021	Nov.2021	新加坡
36	United Nations Educational Scientific and Cultural Organization (UNESCO)	Recommendation on the Ethics of Artificial Intelligence	2021	Nov.2021	国际组织
37	Federal Trade Commission (FTC)	Final Amendments to Safeguards Rule become effective	2022	Jan.2022	美国
38	Bank of England, Financial Conduct Authority	Artificial Intelligence Public-Private Forum Final Report	2022	Feb.2022	英国
39	Consumer Financial Protection Bureau (CFPB)	Circular on adverse action requirements for complex algorithms	2022	May.2022	美国
40	Department for Business, Energy & Industrial Strategy, Department for Digital, Culture, Media & Sport, Department for Science, Innovation & Technology, Office for Artificial Intelligence	Policy paper: Establishing a pro-innovation approach to regulating AI	2022	Jul.2022	英国
41	Federal Trade Commission (FTC)	ANPR on commercial surveillance and data security	2022	Aug.2022	美国
42	the European Commission	Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)	2022	Sep.2022	欧盟
43	White House	Blueprint for AI Bill of Rights	2022	Oct.2022	美国
44	Bank of England	DP5/22 - Artificial Intelligence and Machine Learning (Discussion Paper)	2022	Oct.2022	英国
45	Bank of England and Financial Conduct Authority	Research Note: Machine learning in UK financial services 2022	2022	Oct.2022	英国
46	UK's Information Commissioner's Office	Explaining decisions made with AI	2022	Oct.2022	英国
47	中国银行业协会	《人工智能模型风险管理框架》	2022	Oct.2022	中国
48	California	CPPA/CPRA became effective	2023	Jan. 2023	美国
49	Virginia	Consumer Data Protection Act (VCDPA) became effective	2023	Jan. 2023	美国
50	National Institute of Standards and Technology (NIST)	Voluntary AI Risk Management Framework	2023	Jan. 2023	美国
51	Federal Trade Commission (FTC)	Blogs on advertising AI products/services and AI & UDAP	2023	Feb/Mar. 2023	美国
52	International Organization for Standardization (ISO) /International Electrotechnical Commission (IEC)	ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management	2023	Feb. 2023	国际组织

序号	立法机构	文件名称	年度	具体时间	国别/区域
53	Department for Science, Innovation & Technology	AI regulation: a pro-innovation approach	2023	Mar.2023	英国
54	UK's Information Commissioner's Office	Guidance on AI and data protection	2023	Mar.2023	英国
55	Securities and Exchange Commission (SEC)	Proposal to amend Regulation S-P	2023	Mar. 2023	美国
56	The Consumer Financial Protection Bureau (CFPB), the Department of Justice Civil Rights Division (DOJ), the Equal Employment Opportunity Commission (EEOC), and the Federal Trade Commission (FTC)	Interagency Statement on supervision and enforcement of "automated systems"	2023	Apr. 2023	美国
57	Department of Commerce (DOC)	NTIA Request for Comment on AI assurance and trust	2023	Apr. 2023	美国
58	Montana	Enacted the Consumer Data Privacy Act	2023	May. 2023	美国
59	Federal Trade Commission (FTC)	Policy Statement on Misuse of Biometric Info	2023	May. 2023	美国
60	Congress	Multiple hearings on AI, oversight, and intellectual property	2023	May. 2023	美国
61	Interagency: Office of the Comptroller of the Currency (OCC), US Treasury; Board of Governors of the Federal Reserve System (Board); Federal Deposit Insurance Corporation (FDIC); National Credit Union Administration (NCUA); Consumer Financial Protection Bureau (CFPB); and Federal Housing Finance Agency (FHFA)	NPR on amendments to FIRREA regarding automated valuation models	2023	Jun. 2023	美国
62	New York City	Automated Employment Decision Tool law becomes effective	2023	Jul. 2023	美国
63	Securities and Exchange Commission (SEC)	NPR on Conflicts of Interest using predictive data analytics	2023	Jul. 2023	美国
64	Colorado	Colorado Privacy Act becomes effective	2023	Jul. 2023	美国
65	Connecticut	Connecticut Data Privacy Act becomes effective	2023	Jul. 2023	美国
66	国家互联网信息办公室, 国家发展和改革委员会, 教育部, 科学技术部, 工业和信息化部, 公安部, 国家广播电视总局	《生成式人工智能服务管理暂行办法》	2023	Jul. 2023	中国
67	Connecticut	Act concerning Artificial intelligence, automated decision-making and personal data privacy	2023	Jul. 2023	美国



序号	立法机构	文件名称	年度	具体时间	国别/区域
68	全国信息安全标准化技术委员会	关于发布《网络安全标准实践指南——生成式人工智能服务内容标识方法》的通知	2023	Aug. 2023	中国
69	国家市场监督管理总局、国家标准化管理委员会	《信息安全技术机器学习算法安全评估规范》(GB/T 42888-2023)	2023	Aug. 2023	中国
70	White House	Launch of AI Cyber Challenge	2023	Aug. 2023	美国
71	中国信息通信研究院	《面向行业的大规模预训练模型技术和应用评估方法 第1部分：金融大模型》	2023	Sep. 2023	中国
72	Cybersecurity and Infrastructure Security Agency (CISA), with NSA/FBI	cybersecurity info sheet on deepfakes	2023	Sep. 2023	美国
73	White House	Fact Sheet on voluntary commitments from AI companies on AI risk management	2023	Sep. 2023	美国
74	Consumer Financial Protection Bureau (CFPB)	Guidance on Credit Denials by Lenders Using Artificial Intelligence	2023	Sep. 2023	美国
75	White House	Executive Order on Safe, Secure and Trustworthy AI	2023	Oct. 2023	美国
76	the Group of Seven (G7)	G7 Leaders' Statement on the Hiroshima AI Process the Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems and the Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems	2023	Oct. 2023	国际组织
77	Financial Conduct Authority (FCA)	FS23/6: Artificial Intelligence and Machine Learning (Feedback Statement)	2023	Oct. 2023	英国
78	国家互联网信息办公室	《全球人工智能治理倡议》	2023	Oct. 2023	中国
79	首届全球人工智能 (AI) 安全峰会	Bletchley Declaration 《布莱切利宣言》	2023	Nov. 2023	国际组织
80	National Institute of Standards and Technology (NIST)	Request for Information (RFI) to assist with implementation of its responsibilities under the recent Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence	2023	Dec. 2023	美国
81	The Institute of Internal Auditors	Artificial Intelligence Auditing Framework	2023	-	国际组织
82	Bank for International Settlements (BIS)	Artificial intelligence in central banking	2024	Jan. 2024	国际组织
83	White House	Financial Services Committee bipartisan working group on AI	2024	Jan. 2024	美国

序号	立法机构	文件名称	年度	具体时间	国别/区域
84	New York State Department of Finance Services (NYDFS) 纽约州金融服务部	Proposed guidance on AI in the insurance industry	2024	Jan.2024	美国
85	Commodity Futures Trading Commission (CFTC)	RFC on AI uses/risks	2024	Jan.2024	美国
86	White House	Fact Sheet on AI Actions related to EO	2024	Jan.2024	美国
87	Department for Science, Innovation & Technology 英国政府	A pro-innovation approach to AI regulation: government response	2024	Feb.2024	英国
88	Department for Science, Innovation & Technology, Prime Minister's Office	Policy paper: The UK Science and Technology Framework	2024	Feb.2024	英国
89	Utah	Artificial Intelligence Policy Act	2024	Mar.2024	美国
90	Hong Kong Monetary Authority (HKMA)	Thematic Review of Transaction Monitoring Systems and Use of Artificial Intelligence	2024	Apr.2024	香港
91	Bank for International Settlements (BIS)	Project Raven: using AI to assess financial system's cyber security and resilience	2024	Apr.2024	国际组织
92	Financial Conduct Authority (FCA)	AI Update	2024	Apr.2024	英国
93	Council of Europe	Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law	2024	May.2024	欧洲
94	Council of Europe	the EU Artificial Intelligence Act (draft version in Dec.2023)	2024	May.2024	欧洲
95	Bank for International Settlements (BIS)	Generative artificial intelligence and cyber security in central banking	2024	May.2024	国际组织
96	The Bipartisan Senate AI Working Group	Driving U.S. Innovation in Artificial Intelligence	2024	May.2024	美国
97	National Institute of Standards and Technology (NIST)	Draft-Guidance: Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence (GenAI) Profile (NIST AI 600-1)	2024	May.2024	美国
98	National Institute of Standards and Technology (NIST)	Draft-Guidance: Secure Software Development Practices for GenAI and Dual-Use Foundation Models (NIST Special Publication (SP) 800-218A)	2024	May.2024	美国
99	National Institute of Standards and Technology (NIST)	Draft-Report: Reducing Risks Posed by Synthetic Content (NIST AI 100-4)	2024	May.2024	美国

序号	立法机构	文件名称	年度	具体时间	国别/区域
100	National Institute of Standards and Technology (NIST)	Draft-Plan: Global Collaboration on AI Standards (NIST AI 100-5)	2024	May.2024	美国
101	Colorado	Colorado Artificial Intelligence Act (CAIA)	2024	Jun.2024	美国
102	Bank for International Settlements (BIS)	Intelligent financial system: how AI is transforming finance	2024	Jun.2024	国际组织
103	European Central Bank	Digital innovation and banking regulation	2024	Jun.2024	欧洲
104	Financial Industry Regulatory Authority (FINRA)	Regulatory Notice 24-09: FINRA Reminds Members of Regulatory Obligations When Using Generative Artificial Intelligence and Large Language Models	2024	Jun.2024	美国
105	the European Commission	A targeted consultation and a workshop series to seek input from stakeholders on the use of artificial intelligence (AI) in finance	2024	Jun.2024	欧盟
106	Department of the Treasury	Request for Information on Uses, Opportunities, and Risks of Artificial Intelligence in the Financial Services Sector	2024	Jun.2024	美国
107	2024世界人工智能大会暨人工智能全球治理高级别会议	《人工智能全球治理上海宣言》	2024	Jul.2024	中国
108	欧洲保险与职业养老金局	关于适用于保险行业人工智能系统的监管框架的情况说明书	2024	Jul.2024	欧洲

## 附件二：各地区/国家人工智能相关监管机构清单

机构名称	机构名称 (中文)	国家/地区
Organisation for Economic Cooperation and Development (OECD)	世界经济合作组织	国际组织
G20	二十国集团	国际组织
the Group of Seven (G7)	七国集团	国际组织
United Nations Educational Scientific and Cultural Organization (UNESCO)	联合国教科文组织	国际组织
International Organization for Standardization (ISO)	国际标准化组织	国际组织
International Electrotechnical Commission (IEC)	国际电工委员会	国际组织
International Organization of Securities Commissions (IOSCO)	国际证券事务监察委员会组织	国际组织
Bank for International Settlements (BIS)	国际清算银行	国际组织
The Institute of Internal Auditors (IIA)	国际内部审计师协会	国际组织
Council of Europe	欧洲理事会	欧盟/欧洲
the European Commission	欧盟委员会	欧盟/欧洲
Independent High-level Expert Group on AI (set up by the European Commission)	欧盟委员会	欧盟/欧洲
European Banking Authority	欧洲银行管理局	欧盟/欧洲
European Central Bank	欧洲中央银行	欧盟/欧洲
European Insurance and Occupational Pensions Authority	欧洲保险与职业养老金局	欧盟/欧洲
Bank of England	英格兰银行	英国
Financial Conduct Authority (FCA)	金融行为监管局	英国
UK's Information Commissioner's Office	信息专员办公室	英国
Department for Science, Innovation and Technology	科学、创新和技术部	英国
Office for Artificial Intelligence	人工智能办公室	英国
Department for Digital, Culture, Media & Sport	英国数字、文化、媒体和体育部	英国
Department for Business, Energy & Industrial Strategy	商业、能源和工业战略部	英国
White House	白宫	美国
Congress	议会	美国
The Bipartisan Senate AI Working Group	两党参议院人工智能工作组	美国
US Treasury	财政部	美国
Department of Commerce (DOC)	商务部	美国
Federal Trade Commission (FTC)	联邦贸易委员会	美国
Securities and Exchange Commission (SEC)	证券交易委员会	美国
Consumer Financial Protection Bureau (CFPB)	消费者金融保护局	美国
National Institute of Standards and Technology (NIST)	美国国家标准与技术研究院	美国
the Department of Justice Civil Rights Division (DOJ)	司法部民权司	美国
the Equal Employment Opportunity Commission (EEOC)	平等就业机会委员会	美国
Office of the Comptroller of the Currency (OCC)	货币监理署	美国
Board of Governors of the Federal Reserve System (Board)	联邦储备系统理事会	美国
Federal Deposit Insurance Corporation (FDIC)	联邦存款保险公司	美国
National Credit Union Administration (NCUA)	全国信用合作社管理局	美国



机构名称	机构名称 (中文)	国家/地区
Federal Housing Finance Agency (FHFA)	联邦住房金融局	美国
Cybersecurity and Infrastructure Security Agency (CISA)	网络安全和基础设施安全局	美国
Commodity Futures Trading Commission (CFTC)	商品期货交易委员会	美国
National Telecommunications and Information Administration (NTIA)	国家电信和信息管理局	美国
National Association of Insurance Commissioners	保险专员协会	美国
New York State Department of Finance Services (NYDFS)	纽约州金融服务部	美国
Financial Industry Regulatory Authority (FINRA)	金融业监管局	美国
国务院	国务院	中国
国家新一代人工智能治理专业委员会	国家新一代人工智能治理专业委员会	中国
国家发展和改革委员会	国家发展和改革委员会	中国
教育部	教育部	中国
科学技术部	科学技术部	中国
工业和信息化部	工业和信息化部	中国
公安部	公安部	中国
国家广播电视总局	国家广播电视总局	中国
全国信息安全标准化技术委员会	全国信息安全标准化技术委员会	中国
国家互联网信息办公室	国家互联网信息办公室	中国
中国信息通信研究院	中国信息通信研究院	中国
Federal Financial Supervisory Authority of Germany (BaFin)	联邦金融监管局	德国
Monetary Authority of Singapore (MAS)	新加坡金融管理局	新加坡
Financial Sector Supervisory Commission of Luxembourg (CSSF)	卢森堡金融部门监督委员会	卢森堡
Netherlands Bank	荷兰银行	荷兰
French Prudential Supervision and Resolution Authority (ACPR)	法国审慎监管和处置机构	法国
Hong Kong Monetary Authority (HKMA)	香港金融管理局	香港

## 联系我们

---

### 李砾

金融研究中心合伙人

毕马威中国

电话: +86 139 1033 7443

邮箱: raymond.li@kpmg.com

### 靳蓓

金融研究中心副总监

毕马威中国

电话: +86 136 1115 0725

邮箱: catherine.jin@kpmg.com

### 赵一诺

金融研究中心副总监

毕马威中国

电话: +86 139 1170 3295

邮箱: nicholas.y.zhao@kpmg.com

[kpmg.com/cn/socialmedia](https://kpmg.com/cn/socialmedia)



如需获取毕马威中国各办公室信息，请扫描二维码或登陆我们的网站：  
<https://home.kpmg/cn/zh/home/about/offices.html>

所载资料仅供一般参考用，并非针对任何个人或团体的个别情况而提供。虽然本所已致力提供准确和及时的资料，但本所不能保证这些资料在阁下收取时或日后仍然准确。任何人士不应在没有详细考虑相关的情况及获取适当的专业意见下依据所载资料行事。

© 2024 毕马威华振会计师事务所 (特殊普通合伙) — 中国合伙制会计师事务所，是与毕马威国际有限公司 (英国私营担保有限公司) 相关联的独立成员所全球组织中的成员。版权所有，不得转载。在中国印刷。

毕马威的名称和标识均为毕马威全球组织中的独立成员所经许可后使用的商标。

二零二四年九月出版