



2023 年 10 月  
人工智能月刊  
(2023.10.1-2023.10.31)



植德律师事务所元宇宙与数字经济行业委员会 AIGC 研究小组

中国科学院大学经济与管理学院孙毅教授课题组

(排名不分先后)

## 导读

### 最新行业动态

1. Stability AI 发布 Stable LM 3B 语言模型
2. Nature: ChatGPT 和其他人工智能工具将如何颠覆科学出版
3. Adobe 发布了新一代 Firefly 模型
4. Gartner: 到 2026 年, 超过 80% 的企业将使用生成式 AI api 或部署生成式 AI 应用程序
5. 谷歌: 保护使用生成式 AI 的消费者
6. 斯坦福大学发布《基础模型透明度指数》
7. IBM NorthPole: 新的芯片架构指向更快、更节能的人工智能
8. 英伟达发布 AI 系统 Eureka, 可训练机器人完成复杂任务
9. 阿里达摩院发布首个遥感 AI 大模型
10. 讯飞星火 V3.0 发布
11. OpenAI 正在组建团队应对 AI 的灾难性风险
12. 高通发布 PC 芯片, 专为 AI 打造
13. 波士顿动力最新研究将 AI 模型集成到机器人
14. 百度文心一言开启收费模式

### 最新法律与监管动态

1. 科技部、工信部等 10 部委联合发布《科技伦理审查办法（试行）》
2. 《生成式人工智能服务 安全基本要求》（征求意见稿）发布, 向社会征求意见
3. 工业和信息化部印发《人形机器人创新发展指导意见》

4. 欧盟委员会发布最新人工智能示范合同条款草案
5. 加利福尼亚州州长签署《删除法案》
6. 美国总统拜登发布关于安全、可靠和值得信赖的人工智能的行政命令
7. 七国集团宣布同意遵守《人工智能国际指导原则》和《开发先进人工智能系统组织的国际行为准则》
8. 国家金融监督管理总局提示广大金融消费者警惕利用 AI 新型技术实施诈骗的风险
9. 习近平主席在第三届“一带一路”国际合作高峰论坛开幕式提出《全球人工智能治理倡议》
10. FTC 发布消费者对人工智能表示担忧的报告
11. 乌克兰发布《人工智能 (AI) 监管路线图》
12. 印度电子和信息技术部发布《印度人工智能 2023 计划》报告
13. 美国、新加坡就人工智能发表联合声明
14. 英国人工智能安全峰会发布《布莱切利宣言》，包括中国在内的 28 国及欧盟共同签署
15. 环球音乐起诉人工智能公司 Anthropic 侵犯版权

## 一、最新行业动态

### 1. Stability AI 发布 Stable LM 3B 语言模型

发布日期：2023 年 10 月 2 日

来源：Stability AI

链接：

- <https://stability.ai/blog/stable-lm-3b-sustainable-high-performance-language-models-smart-devices>

摘要：

据称 Stable LM 3B 有 30 亿个参数(业界通常使用的 70 到 700 亿个参数)，相对来说是一种紧凑的语言模型，设计用于便携式数字设备，如手持设备和笔记本电脑。与大型模型不同，模型需要更少的资源和更低的运营成本，使大多数用户都可以轻松访问它们。这不仅使它们更便宜，而且还使它们更环保，因为它们消耗的电力要少得多。同时，Stable LM 3B 具有很强的竞争力——它超越了以前最先进的 3B 参数语言模型，甚至在 7B 参数尺度上超越了一些最好的开源语言模型。

### 2. Nature: ChatGPT 和其他人工智能工具将如何颠覆科学出版

发布日期：2023 年 10 月 10 日

来源：Nature

链接：

- <https://www.nature.com/articles/d41586-023-03144-w>

摘要：

尽管《自然》杂志的一项调查显示，经常使用大语言的科学家仍占少数，但许多人预计，生成式人工智能工具将成为撰写手稿、同行评议报告和申请资助的常规助手。这些只是人工智能可以改变科学传播和出版的一些方式。科学出版商已经在科学搜索工具以及编辑和快速总结论文中试用生成式人工智能。许多研究人员认为，非英语母语者可以从这些工具中获益最多。一些人认为，生成式人工智能是科学家重新思考他们如何询问和总结实验结果的一种方式——他们可以使用大语言模型来完成大部分工作，这意味着更少的时间写论文，更多的时间做实验。

加州大学伯克利分校的计算生物学家 Michael Eisen 预测，生成式人工智能工具甚至可以从根本上改变科学论文的性质。但是出版商担心，使用人工智能的增加可能会导致更多质量差或错误的手稿——可能会出现大量人工智能辅助的赝品。

### 图灵财经短评

人工智能，特别是生成式人工智能已经对社会各个方面产生深刻影响。学术界也不例外，目前来看，不论出版集团还是研究人员对人工智能的看法仍然褒贬不一。从 Chatgpt 横空出世以来，在学术界就引发了不小震动，不少高校明令禁止使用这类 AI 工具，比如香港大学向师生发出内部邮件，表明禁止在港大所有课堂、作业和评估中使用 ChatGPT 或其他 AI 工具，许多世界顶尖高校表明同样态度。随着大语言模型的发展以及国家的肯定态度，似乎对此类工具有了新的态度转变，除了上述报道，Nature 还做过一项针对全球的博士后进行的调查，结果表明有三分之一的受访者正在使用 AI 聊天机器人来帮助自己修改文字、生成或编辑代码、整理相关领域的文献等工作。结合日前在科技部科技监督与诚信建设司的指导和帮助下，中国科学技术信息研究所牵头联合施普林格·自然（等三家国际出版集团出版的《学术出版中 AIGC 使用边界指南》，可以看出，如何加强 AIGC 在学术出版过程中的规范化使用应是未来的发展方向。此外，学术研究人员在利用工具提高效率的同时，更应该坚守学术道德。

### 3. Adobe 发布了新一代 Firefly 模型

发布日期：2023 年 10 月 10 日

来源：Adobe

链接：

- <https://news.adobe.com/news/news-details/2023/Adobe-Releases-Next-Generation-of-Firefly-Models/default.aspx>

摘要：

在全球最大的创意大会 Adobe MAX 上，Adobe 发布了 Adobe Firefly Image 2 Model、Adobe Firefly Vector Model 和 Adobe Firefly Design Model 三款模型。Firefly Image 2 是下一代生成式人工智能成像，为用户带来了一个新的水平的创意控制和质量。Firefly Vector 是世界上第一个专注于生成矢量图形的生成式 AI 模型，也是第一个生成“人类品质”矢量和图案输出的生成式 AI 模型，将 Adobe 的矢量图形和生成式 AI 专业知识直接带入 Adobe Illustrator 的工作流程中。Firefly Design 通过新的“文本到模板”功能，在 Adobe Express 中即时生成令人惊叹的高质量模板设计。

### 4. Gartner：到 2026 年，超过 80% 的企业将使用生成式 AI api 或部署生成式 AI 应用程序

发布日期：2023 年 10 月 11 日

来源：Gartner

链接:

- <https://www.gartner.com/en/newsroom/press-releases/2023-10-11-gartner-says-more-than-80-percent-of-enterprises-will-have-used-generative-ai-apis-or-deployed-generative-ai-enabled-applications-by-2026>

摘要:

根据 Gartner 公司的数据,到 2026 年,超过 80%的企业将使用生成式人工智能 GenAI 应用程序编程接口 api 或模型,或在生产环境中部署支持 GenAI 的应用程序,而 2023 年这一比例还不到 5%。Gartner 杰出副总裁分析师 Arun Chandrasekaran 表示:“在医疗保健、生命科学、法律、金融服务和公共部门等许多行业,对生成式人工智能的需求正在增加。”

Gartner 预计在十年内对组织产生巨大影响的三项创新包括支持 Genai 的应用程序、基础模型和 AI 信任、风险和安全管理。

## 5. 谷歌:保护使用生成式 AI 的消费者

发布日期:2023 年 10 月 13 日

来源:谷歌

链接:

- <https://cloud.google.com/blog/products/ai-machine-learning/protecting-customers-with-generative-ai-indemnification>

摘要:

谷歌称自己的客户如果在版权方面受到挑战,公司将承担潜在的法律风险。谷歌将从两个方面的赔偿来让用户在使用我们的生成式人工智能产品时更加放心,第一个问题涉及谷歌对训练数据的使用,而第二个问题专门涉及基础模型的生成输出。

此外,作为客户也应当发挥作用。例如此赔偿仅适用于没有试图故意创建或使用生成的输出来侵犯他人的权利;以及在使用现有的和新兴的工具,例如引用来源以帮助负责任地使用生成的输出。

### 图灵财经短评

生成式人工智能表现出强大的生产力,同时任何人只需简单了解便可熟练使用,但这也带来了诸多挑战。今年年初,全球知名图片提供商华盖创意起诉热门人工智能绘画工具 Stable Diffusion 的开发者 Stability AI,称其侵犯了版权,此外近日北京互联网法院依法公开审理了一起 AI 文生图著作权案,关于版权的侵权案件已经引发了多方关注。大语言模型的训练数据大多数来源于网页数据,如果没有标注,用户大多数情况下是不了解生成结果的来源。谷歌的这项措施既是应对监管的行动,在各大厂商大模型纷纷推出的背景下,也是吸引消费者的一种手

段。除了政府对数字版权的立法规定外，大模型更应该增加其透明度，对数据来源应作出详尽的说明，减少此类事件发生。

## 6. 斯坦福大学发布《基础模型透明度指数》

发布日期：2023 年 10 月 18 日

来源：斯坦福大学

链接：

- <https://hai.stanford.edu/news/introducing-foundation-model-transparency-index>

摘要：

来自斯坦福大学、麻省理工学院和普林斯顿大学的多学科团队为评估透明度，设计了一个名为基础模型透明度指数的评分系统。该模型评估了透明度的 100 个不同方面，从公司如何建立基础模型，如何运作，以及如何在下游使用。团队使用 100 分指数对 10 家主要公司进行评分时，发现最高分在 47 到 54 之间，而最低值只有 12 分。

大约三分之一的指标与基础模型开发人员如何构建模型有关，包括关于训练数据的信息，用于创建模型的劳动，以及所涉及的计算资源。另外三分之一与模型本身有关，包括其能力、可信度、风险以及这些风险的缓解。最后三分之一涉及如何在下游使用模型，包括公司关于模型分发、用户数据保护和模型行为的政策，以及公司是否为受影响的个人提供反馈或纠正的机会。

根据得分，排名第一的是 Meta 的 Llama 2，但分数仅为 54；GPT-4 分数仅仅为 48，排名第三；最后一名是 Amazon 的 Titan Text，仅取得 12 分。

### 图灵财经短评

从目前大模型的发展来看，不管像 ChatGPT 这样的顶尖的算法，还是其他厂商的模型也会存在一个共性问题，那就是模型的质量取决于用于训练的数据质量。但导致不良模型可能有多种原因，除了输入的数据存在问题，还存在数据清理不到位、模型的参数设置不当等因素。因此，一旦模型出现错误却不容易找到其根本原因。这就是由模型的不透明引发的“信任危机”。

总体上说，大模型的不透明性主要有几个方面，一是一些算法的不可解释性；二是训练数据的选择及其来源的不明确；三是模型持续迭代过程不透明。从公司利益出发，大模型不透明可能是开发公司为了避免法律的约束以及保护自身的商业机密、甚至创造行业壁垒。种种原因导致很多模型是处于“保密”状态，达不到可信任的程度，斯坦福的数据显示没有一个模型能够达到及格水平，因此加强监管的同时，更需要一份权威的评估标准解决这个问题。

## 7. IBM NorthPole: 新的芯片架构指向更快、更节能的人工智能

发布日期: 2023 年 10 月 20 日

来源: IBM

链接:

- <https://research.ibm.com/blog/northpole-ibm-ai-chip>

摘要:

加州 IBM 研究院(IBM Research)实验室研发的一款新型芯片声称有可能颠覆人工智能的有效使用方式和领域。计算机芯片基本上遵循相同的基本结构,即处理单元和存储待处理信息的存储器分开存储,但这种结构在芯片内的存储器、处理器和任何其他设备之间不断地来回移动数据需要花费时间和精力。

研究人员 Modha 透露, NorthPole 是芯片架构上的一个突破,它在能源、空间和时间效率方面带来了巨大的改善。使用 ResNet-50 模型作为基准, NorthPole 比普通的 12 纳米 gpu 和 14 纳米 cpu 效率高得多。根据 Modha 的说法,在 ResNet-50 上, NorthPole 的性能优于所有主流架构,即使是那些使用更先进技术工艺的架构,比如使用 4nm 工艺实现的 GPU。

## 8. 英伟达发布 AI 系统 Eureka, 可训练机器人完成复杂任务

发布日期: 2023 年 10 月 20 日

来源: 英伟达

链接:

- <https://blogs.nvidia.com/blog/2023/10/20/eureka-robotics-research/>

摘要:

英伟达研究中心宣称开发了一种新的人工智能代理,可以教机器人复杂的技能,已经训练出一只机器人手来执行快速旋转笔的技巧。除此之外,还能教机器人打开抽屉和橱柜,抛球和接球,操纵剪刀,以及其他任务,这是由于 Eureka 可以自动编写奖励算法来训练机器人。

Eureka 生成的奖励程序—使机器人能够进行试错学习—在 80% 以上的任务中表现优于人类编写的专家程序。这使得机器人的平均性能提高了 50% 以上。Eureka 利用 GPT-4 LLM 和生成式人工智能编写软件代码,奖励机器人进行强化学习。它不需要特定于任务的提示或预定义的奖励模板——并且很容易结合人类反馈来修改结果的奖励,从而更准确地与开发人员的想法保持一致。

## 9. 阿里达摩院发布首个遥感 AI 大模型

发布日期: 2023 年 10 月 20 日



来源：达摩院

链接：

- [https://mp.weixin.qq.com/s/Ni\\_gld91PNiZ0jPJ0dpxHA](https://mp.weixin.qq.com/s/Ni_gld91PNiZ0jPJ0dpxHA)

摘要：

遥感技术在城市运营、耕地保护、应急救援等国计民生中应用甚广，遥感 AI 则可大幅提升数据的利用深度，提供更智能的分析应用，如结合卫星照片与历史气象情况，“算”出某一块农田里作物的长势状况，让种地不再被动，而是更主动地“看天吃饭”。阿里达摩院发布业内首个遥感 AI 大模型（AIE-SEG），率先在遥感领域实现了图像分割的任务统一，一个模型即可实现“万物零样本”的快速提取，可识别农田、水域、建筑物等近百种遥感地物分类，还能根据用户的交互式反馈自动调优识别结果。达摩院称该模型在一些特定场景下，对比传统的遥感模型，实例提取的准确率可提升 25%，变化检测的准确率可提升 30%。基于这些能力，遥感 AI 大模型提供了“开箱即用”的 API 调用服务，用户可以根据需求定制不同的遥感 AI 解译功能，如水体提取、耕地变化监测、光伏识别等。

## 10. 讯飞星火 V3.0 发布

发布日期：2023 年 10 月 24 日

来源：科大讯飞

链接：

- <https://mp.weixin.qq.com/s/-fZT0jyeXosjqTTaFfyCdw>

摘要：

在第六届世界声博会暨 2023 科大讯飞全球 1024 开发者节，科大讯飞发布讯飞星火认知大模型 V3.0，七大能力持续提升，整体超越 ChatGPT，医疗领域六大核心能力超越 GPT-4。星火 V3.0 升级启发式对话、AI 人设功能，能够打造更个性化的 AI 助手。不过，科大讯飞董事长刘庆峰也提到，国产大模型必须正视和 GPT-4 的差距。“国产大模型在复杂知识推理、小样本快速学习、超长文本处理、跨模态统一理解上距 GPT4 还有差距。”他表示，对标 GPT-4 的更大参数规模的星火大模型正式启动训练，2024 年上半年将实现对标 GPT-4。

## 11. OpenAI 正在组建团队应对 AI 的灾难性风险

发布日期：2023 年 10 月 26 日

来源：OpenAI

链接：

- <https://openai.com/blog/frontier-risk-and-preparedness>

**摘要：**

OpenAI 声称将认真对待与人工智能相关的各种安全风险，并在今年 7 月与其他领先的人工智能实验室一起做出了一系列自愿承诺，以促进人工智能的安全、保障和信任。OpenAI 认为前沿人工智能构成了越来越严重的风险，管理前沿人工智能带来的灾难性风险需要解决以下问题：无论是现在还是将来，前沿人工智能系统被误用会有多危险？我们如何建立一个强大的框架来监测、评估、预测和保护前沿人工智能系统的危险能力？如果我们的前沿人工智能模型权重被窃取，恶意行为者会如何利用它们？

OpenAI 正在组建一个新团队，该团队将帮助跟踪、评估、预测和防范多个类别的灾难性风险。

**12. 高通发布 PC 芯片，专为 AI 打造**

发布日期：2023 年 10 月 26 日

来源：高通

**链接：**

- <https://www.qualcomm.com/news/onq/2023/10/rethink-whats-possible-with-new-snapdragon-x-elite-platform>

**摘要：**

高通近日发布一款 PC 芯片，名为 Snapdragon X Elite。高通声称 Snapdragon X Elite 是一款突破性的 4nm 芯片，将提供了超越同类产品的极致系统级性能、效率和更智能的用户体验。

Snapdragon X Elite 可以以每秒 30 个 token 的速度高效处理 70 亿个参数大语言模型，比如 Llama 2，这将大幅提升人工智能处理速度并提供低延迟、先进的多任务处理。此外，还将支持更大、更复杂的模型，比如那些在设备上运行超过 130 亿个参数的模型，而无需利用云资源。对比英特尔、苹果等产品，高通强调 Snapdragon X Elite 具有更低的功耗，AI 处理速度超过竞品的 4.5 倍。

**13. 波士顿动力最新研究将 AI 模型集成到机器人**

发布日期：2023 年 10 月 27 日

来源：波士顿动力

**链接：**

- <https://bostondynamics.com/blog/robots-that-can-chat/>

**摘要：**

由于人工智能，特别是“生成式人工智能”领域的发展非常迅速。因此波士

顿研究人员想探索这些模型是如何工作的，以及它们将如何影响机器人技术的发展，他们将诸如 ChatGpt 等大语言模型应用于机器人 Spot，使之变成了能对话、有自主决策能力的导游。

波士顿动力官网显示 Spot 能够根据文字、语音提示与人类进行交谈，同时提供了视觉问答功能，可以分析摄像头拍摄的画面以及自动生成图像说明。此外，波士顿动力还公布了一些相应的技术细节。

#### 14. 百度文心一言开启收费模式

发布日期：2023 年 10 月 31 日

来源：文心一言

链接：

● <https://yiyan.baidu.com/>

摘要：

据文心一言官网显示，在 10 月 31 日，其更新了《文心一言会员服务协议》，意味着率先开启国内大语言收费模式。根据服务协议，目前文心一言提供会员和联合会员两种服务，用户与非会员权益上的区别，体现在模型能力、图片生成、插件权益上，如图片生成功能，会员可生成的图片分辨率 1024×1024，在网页端可一次生成多图。非会员可生成的图片分辨率 512×512，并且不可一次生成多图。此外，免费用户只能够使用文心大模型 3.0 版本，最新发布的 4.0 版本只有开通会员获取权限。

## 二、最新法律与监管动态

### 1. 科技部、工信部等 10 部委联合发布《科技伦理审查办法（试行）》

发布日期：2023 年 10 月 8 日

来源：科技部

链接：

- [https://www.most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2023/202310/t20231008\\_188309.html](https://www.most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2023/202310/t20231008_188309.html)

摘要：

近日科技部、教育部、工业和信息化部、国家卫健委、中国科学院等十部门联合发布《科技伦理审查办法（试行）》（以下简称“《审查办法》”），系继 2023 年 3 月中办、国办印发《关于加强科技伦理治理的意见》，对科技伦理治理的基本模式、监管框架、制度体系等作出顶层设计后，我国科技治理领域又一重磅文件。

#### 植德短评

《审查办法》全文主要对科技伦理审查对象、审查主体、审查程序和监督管理作出了全面阐述。

《审查办法》明确开展以下科技活动应进行科技伦理审查：涉及以人为研究参与者的科技活动，包括以人为测试、调查、观察等研究活动的对象，以及利用人类生物样本、个人信息数据等的科技活动；涉及实验动物的科技活动；不直接涉及人或实验动物，但可能在生命健康、生态环境、公共秩序、可持续发展等方面带来伦理风险挑战的科技活动；依据法律、行政法规和国家有关规定需进行科技伦理审查的其他科技活动。

《审查办法》从审查主体、审查程序和监督管理等方面，提出了具体举措。具体而言，高等学校、科研机构、医疗卫生机构、企业等是本单位科技伦理审查管理的责任主体。从事生命科学、医学、人工智能等科技活动的单位，研究内容涉及科技伦理敏感领域的，应设立科技伦理（审查）委员会。

审查程序则包含了申请与受理、一般程序、简易程序、专家复核程序、应急程序。《审查办法》提到，科技伦理（审查）委员会应对审查批准的科技活动开展伦理跟踪审查，必要时可作出暂停或终止科技活动等决定。跟踪审查间隔一般不超过 12 个月。

有关本《审查办法》的详细解读，请见《元宇宙与数字经济系列 | <科技伦理审查办法（试行）>实施后，AIGC 企业应如何应对？》。

([https://mp.weixin.qq.com/s/RxAadC2Qzzk4wyx9Mi2\\_tw](https://mp.weixin.qq.com/s/RxAadC2Qzzk4wyx9Mi2_tw))

## 2. 《生成式人工智能服务 安全基本要求》（征求意见稿）发布，向社会征求意见

发布日期：2023 年 10 月 11 日

来源：全国信息安全标准化技术委员会

链接：

- <https://www.tc260.org.cn/front/postDetail.html?id=20231011143225>

摘要：

2023 年 10 月 11 日全国信息安全标准化技术委员会发布《生成式人工智能服务 安全基本要求》（征求意见稿）（以下简称“《征求意见稿》”），现面向社会公开征求意见，征求意见时间截止于 2023 年 10 月 25 日。

这是国内首个专门面向生成式 AI 安全领域的规范意见稿，也是对 7 月网信办等七部门推出的《生成式人工智能服务管理暂行办法》的支撑。

### 植德短评

《征求意见稿》首次提出生成式 AI 服务提供者需遵循的安全基本要求，包括语料安全、模型安全、安全措施、安全评估等方面。生成式人工智能服务提供者通过备案需按照本文件中各项要求逐条进行安全性评估，并将评估结果以及证明材料在备案时提交。

在语料安全方面，征求意见稿从来源安全、内容安全、标注安全三方面提出了要求。在模型安全方面，征求意见稿从基础模型使用、生成内容安全、服务透明度、内容生成准确性、内容生成可靠性五大方面做出了严格要求。在安全措施方面，征求意见稿从模型适用人群、场合、用途，个人信息处理，手机使用者输入信息用于训练，图片、视频等内容标识，接受公众或使用者的投诉举报，向使用者提供生成内容，模型更新、升级等七大方面提出了要求。在安全评估方面，征求意见稿从评估方法、语料安全评估、生成内容安全评估、内容拒答评估四方面给出了十分具体的参考。

有关本《征求意见稿》的详细解读，请见《元宇宙与数字经济 | 信安标委技术文件<生成式人工智能服务 安全基本要求>（征求意见稿）——生成式人工智能安全评估标准文件解读》。

[（https://mp.weixin.qq.com/s/MeYu-gvPaxeReRc7XhUpow）](https://mp.weixin.qq.com/s/MeYu-gvPaxeReRc7XhUpow)

## 3. 工业和信息化部印发《人形机器人创新发展指导意见》

发布日期：2023 年 11 月 2 日

来源：工信部

链接：

- [https://wap.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2023/art\\_48fe01d562644aedb7ea3f4256df8190.html](https://wap.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2023/art_48fe01d562644aedb7ea3f4256df8190.html)

#### 摘要：

《意见》认为人形机器人集成人工智能、高端制造、新材料等先进技术，有望成为继计算机、智能手机、新能源汽车后的颠覆性产品，将深刻变革人类生产生活方式，重塑全球产业发展格局。

《意见》指出到 2025 年，人形机器人创新体系初步建立，“大脑、小脑、肢体”等一批关键技术取得突破，确保核心部组件安全有效供给。整机产品达到国际先进水平，并实现批量生产，在特种、制造、民生服务等场景得到示范应用，探索形成有效的治理机制和手段。培育 2—3 家有全球影响力的生态型企业和一批专精特新中小企业，打造 2—3 个产业发展集聚区，孕育开拓一批新业务、新模式、新业态。到 2027 年，人形机器人技术创新能力显著提升，形成安全可靠的产业链供应链体系，构建具有国际竞争力的产业生态，综合实力达到世界先进水平。产业加速实现规模化发展，应用场景更加丰富，相关产品深度融入实体经济，成为重要的经济增长新引擎。

此外，《意见》还部署了 5 方面任务：（1）在关键技术突破方面，要打造人形机器人“大脑”和“小脑”、突破“肢体”关键技术、健全技术创新体系。（2）在产品培育方面，要打造整机产品、夯实基础部组件、推动软件创新。（3）在场景拓展方面，实现服务特种领域需求、打造制造业典型场景、加快民生及重点行业推广。（4）在生态营造方面，要培育优质企业、完善创新载体和开源环境、推动产业集聚发展。（5）在支撑能力方面，要健全产业标准体系、提升检验检测和中试验证能力、加强安全治理能力。

#### 植德短评

该《意见》共指导思想、突破关键技术、培育重点产品、拓展场景应用、营造产业生态、强化支撑能力、保障措施七章，为我国发展人形机器人确定了发展目标与前进方向。集中发展人形机器人关键技术，打造人形机器人全生态，培育人形机器人产品应用及产业生态，加快特种领域、制造业、民生等多行业应用场景。并提出措施强化人形机器人技术检测与安全治理能力，在统筹协调、产业政策、人才引进等多方面保障人形机器人与人工智能领域深入发展。该《意见》将极大提升人才投身及科技型企业发展人形机器人产业热情，有效促进我国人形机器人技术进步与产业生态构建。

#### 4. 欧盟委员会发布最新人工智能示范合同条款草案

发布日期：2023 年 10 月 5 日

来源：欧盟委员会官网

链接：

- <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/news/new-version-procurement-clauses-ai-available-supporting-responsible>
- <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai>

#### 摘要：

2023年10月5日，欧盟委员会发布了欧盟人工智能示范合同条款草案（下称“条款草案”），供希望采购外部供应商开发的人工智能系统的公共组织使用。条款草案由人工智能采购社区制定，并经过专家评审。条款草案旨在划定供应商与公共实体之间在以道德、透明和负责任的方式开发人工智能技术方面的责任。

条款草案包含人工智能系统和《人工智能法》所涵盖事项的具体规定，但不包括GDPR等相关法律可能产生的其他义务或要求。欧盟委员会强调，条款草案需要根据不同的合同背景进行定制，因为它们并不包含完整的合同安排，也没有涉及知识产权、付款或责任等方面。

此外，欧盟委员会指出，公共实体与供应商之间的强化数据共享制度现已纳入条款草案，这项制度将数据分为公共组织数据集、供应商数据集和第三方数据集。值得注意的是，欧盟委员会发布了两个版本的条款草案，根据《人工智能法》草案的含义，分别适用于高风险系统和非高风险系统。

欧盟委员会邀请公共组织试用条款草案，并就其使用向人工智能界提供反馈意见。

#### 植德短评

本次发布的人工智能示范合同条款草案经过专家评审，对草案进行部分修改，首先在本次条款草案中增加公共机构与供应商数据共享的相关制度，区分公共机构数据集、供应商数据集和第三方数据集；其次本条款草案增加附件，为协议定制提供参考与现实可能性；另外针对高风险的人工智能系统制定简明版条款草案。

#### 5. 加利福尼亚州州长签署《删除法案》

发布日期：2023年10月10日

来源：加利福尼亚政府官网/加利福尼亚立法数据库

#### 链接：

- <https://www.gov.ca.gov/2023/10/10/governor-newsom-signs-legislation-10-10-23>
- [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill\\_id=202320240SB362](https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB362)

#### 摘要：

10月10日，加利福尼亚州州长 Gavin Newsom 宣布签署参议院第 362 号法案，将与数据经纪人相关的《删除法案》（DELETE Act）纳入法律。

《删除法案》要求加州隐私保护局（the California Privacy Protection Agency, CPPA）在 2026 年 1 月 1 日前建立可访问的删除机制，以允许消费者可以通过单一的、可验证的消费者请求要求维护个人信息的数据经纪人删除其或相关服务提供商或承包商持有的与该消费者相关的任何个人信息。此外，自 2026 年 7 月 1 日起，在消费者提交删除请求以及数据经纪人删除消费者数据后，数据经纪人必须至少每隔 45 天删除一次该消费者的所有个人信息。同时，除非消费者另有要求，否则数据经纪人不得出售或共享消费者新的个人信息。

### 植德短评

《删除法》适用于数据经纪人，其定义是任何“故意收集并向第三方出售与本企业无直接关系的消费者个人信息”的企业。它不包括《公平信用报告法》（FCRA）、《格拉姆-里奇-比利雷法案》（GLBA）、《保险信息和隐私保护法》（HIPAA）所涵盖的实体，也不包括根据《加州民法典》第 1798.146 条对个人信息的处理予以豁免的实体或所涵盖实体的业务关联方（即 HIPAA 所涵盖的实体）。立法者试图“修补”《加州个人信息保护法》中的明显漏洞，即允许消费者要求个别数据经纪人删除直接从消费者处获得的信息，但不要求数据经纪人删除从其他来源获得的个人信息。

## 6. 美国总统拜登发布关于安全、可靠和值得信赖的人工智能的行政命令

发布日期：2023 年 10 月 30 日

来源：美国白宫官网

链接：

- <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>

摘要：

2023 年 10 月 30 日，美国政府拜登发布关于安全、可靠和值得信赖的人工智能的行政命令。该行政命令指示采取以下行动，以保护美国公民免受人工智能潜在风险：（1）强力人工智能系统的开发者与美国政府共享其安全测试结果和其他重要信息；（2）制定标准、工具和测试机制，以帮助确保人工智能系统的安全、可靠和可信；（3）通过制定强有力的生物合成筛选新标准，防范利用人工智能设计危险生物材料的风险；（4）通过制定检测人工智能生成内容和验证官方内容的标准和最佳实践，保护美国人民免受人工智能带来的欺诈和欺骗；（5）建立先进的网络安全计划，开发人工智能工具，发现并修复关键软件中的漏洞；（6）制定一份国家安全备忘录，指导有关人工智能和安全的进一步行动。



另一方面，为了更好地保护美国人民的隐私，行政命令指示采取以下行动：

(1) 优先考虑采用联邦支持，加快开发和使用隐私保护技术，包括使人工智能系统在训练的同时保护训练数据隐私的技术；(2) 加强隐私保护研究和技术发展，如加密工具，为研究协调网络提供资金支持，以推动快速突破和发展；(3) 评估各机构如何收集和使用商业可用信息，并加强联邦机构的隐私指导，以兼顾人工智能；(4) 为联邦机构制定指导方针，以评估隐私保护技术的有效性。

### 图灵财经短评

美国总统拜登签署了美国迄今为止最全面的人工智能监管原则，该行政命令涉及范围广泛，不仅针对美国内部人工智能治理，而且还提及了国际人工智能联合治理。结合日前 G7 领导人就制定人工智能准则达成一致，针对人工智能的监管将达到新的高度。

从去年 10 月的《人工智能权利法案蓝图》为人工智能监管制定了非约束性原则，到今年 7 月 21 日拜登在白宫召集七家领先的人工智能公司，并宣布这些公司已自愿承诺，帮助实现人工智能技术安全、可靠和透明的发展以及 9 月 12 日获得八家企业的第二轮自愿承诺，美国政府已经为这份行政命令做了许多铺垫工作。包括中国、欧盟等国家和地区已经在制定各自的治理方案作出努力，以及即将在英国举行的人工智能安全峰会，表明各国共同推进人工智能监管的共识。

### 植德短评

该行政命令为人工智能安全和保障制定了新标准，旨在保护美国人的隐私，促进公平和公民权利，维护消费者和工人的利益，促进创新和竞争，提升美国在世界各地的领导地位等等。

该行政命令指示领先的人工智能实验室向美国政府通报产生具有潜在国家安全风险的模型的培训运行，指示美国国家标准与技术研究院制定如何对抗性测试人工智能模型的框架，并制定一项计划，利用人工智能来自动查找并修复软件漏洞等措施。针对与人工智能模型相关的隐私、公平性和存在风险等问题，该命令是在世界各地的政策制定者争先恐后地为人工智能制定规则之际为监管制度奠定基础的一次全面尝试。白宫的一份情况说明将该命令描述为“有史以来为保护美国人免受人工智能系统潜在风险而采取的最全面的行动”，试图在利用人工智能的巨大能力和保护“美国人免受人工智能系统的潜在风险”之间划定界限。

## 7. 七国集团宣布同意遵守《人工智能国际指导原则》和《开发先进人工智能系统组织的国际行为准则》

发布日期：2023 年 10 月 30 日

来源：美国白宫官网

链接:

- <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/g7-leaders-statement-on-the-hiroshima-ai-process/>

摘要:

2023年10月30日,七国集团(G7)宣布同意遵守《人工智能国际指导原则》(以下简称“《原则》”)和《开发先进人工智能系统组织的国际行为准则》(以下简称“《行为准则》”)。七国集团指出,将在必要时审查和更新《原则》和行为准则》。

具体而言,《原则》共提出了十一项人工智能相关的原则,其适用于所有人工智能(AI)行为者,为组织开发、部署和使用先进的人工智能系统(如基础模型和生成式人工智能)提供了指导,以提高技术的安全性和可信度。《行为准则》则详细阐述了这些原则,将促进全球人工智能负责任治理。

## 8. 国家金融监督管理总局提示广大金融消费者警惕利用 AI 新型技术实施诈骗的风险

发布日期: 2023年10月13日

来源: 国家金融监督管理总局官网

链接:

- <http://www.cbirc.gov.cn/cn/view/pages/ItemDetail.html?docId=1131422&itemId=915&generaltype=0>

摘要:

10月13日,国家金融监督管理总局发布《关于警惕利用 AI 新型技术实施诈骗的风险提示》(以下简称《提示》),提醒广大金融消费者警惕利用 AI 技术进行的新颖诈骗手段,维护个人及家庭财产安全。

《提示》指出,利用 AI 新型技术实施诈骗主要有“拟声”“换脸”两种手段,即通过模拟他人声音或形象骗取信任,进而诈骗钱财。此类诈骗手段迷惑性、隐蔽性较强,诈骗金额较高。因此,国家金融监督管理总局提示:一、网络渠道“眼见”不一定为实;二、转账汇款务必核验对方身份;三、保护好照片、声音等个人信息。

### 植德短评

随着生成式 AI 的快速发展,“拟声”“换脸”技术日益成熟,不少不法分子发现此处有利可图,便开始利用这些技术实施新型诈骗,利用 AI 新兴技术实施诈骗的案件屡见不鲜,越来越多人因为风险意识的匮乏上当受骗。

此次国家金融监督管理总局的公告,从国家监管层面进行了风险提示,体现

了利用 AI 新型技术诈骗具有迷惑性及隐蔽性强、案件发生频率高的特点。此前，各地公安机关曾破获过数例利用 AI 技术进行刑事犯罪的案例，例如浙江省警方公布的首例团伙利用 ChatGPT 制作虚假视频博取网络流量获利案，济南发生的“AI 换脸诈骗”七秒视频连线被骗 30 万元案等等。

这也提示企业及个人在日后需更好地保护个人信息，提高警惕心，并注意生成式 AI 技术利用时隐藏的刑事法律风险。

## 9. 习近平主席在第三届“一带一路”国际合作高峰论坛开幕式提出《全球人工智能治理倡议》

发布日期：2023 年 10 月 18 日

来源：国家网信办网站

链接：

- [http://www.cac.gov.cn/2023-10/18/c\\_1699291032884978.htm](http://www.cac.gov.cn/2023-10/18/c_1699291032884978.htm)

摘要：

2023 年 10 月 18 日，习近平主席在第三届“一带一路”国际合作高峰论坛开幕式主旨演讲中提出《全球人工智能治理倡议》（以下简称《倡议》），并由国家网信办官网在当天发布。

《倡议》首先提出，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。倡议呼吁全球共同推动人工智能的发展，各国应在人工智能治理中加强信息交流和技术合作，推动构建人类命运共同体。此后，《倡议》提出了十一条细则：明确了人工智能治理应以“以人为本”为理念；坚持“智能向善”的宗旨；坚持“相互尊重、平等互利”的原则；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；逐步建立健全法律和规章制度；坚持公平性和非歧视性原则；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；坚持广泛参与、协商一致、循序渐进的原则；积极发展用于人工智能治理的相关技术开发与应用；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。

### 图灵财经短评

关于人工智能讨论越发频繁，这种技术既是进步更是挑战。如何将人工智能可能出现的负面影响降到最低、如何让人工智能惠及全体人民以及如何使其走向积极发展的一面，这些是当前亟待解决的关键问题。

《倡议》强调安全发展，坚持“智能向善”宗旨，坚持平等互利，共享共治，并支持建立和完善人工智能治理。作为负责任的大国，基于中国实践和中国经验，面对国际提出了中国方案，表明了中方积极推进人工智能全球治理、让人工智能更好赋能世界及各国发展的态度。2023 年，我国已经陆续出台了一系列关

于人工智能治理的政策，未来将持续为构建人类命运共同体贡献中国智慧。

### 植德短评

《倡议》的提出，为全球人工智能治理发出了中国声音、提出中国方案，体现了我国为全人类福祉及人类命运共同体的构建所作的努力和目标。这份《倡议》提出的主要意义在于，推动全球各国共同认识到人工智能治理的必要性，从而共同构建公正、合理、透明的全球人类治理体系，实现全球人工智能的共同发展与进步。

《倡议》虽不具有强制性效力，但其将发挥号召性和引领性的作用，体现了我国的立场及努力方向，特别是对于部分人工智能问题尚未在制度层面上形成具体的规则。《倡议》为正在探索阶段的全球各国的人工智能治理工作提出了“人类命运共同体”的全新治理思路，向全球表明了我国将坚持合作共赢的立场，有利于更好地平衡产业创新与风险管控的关系，协力共同促进人工智能治理。

同时，《倡议》也是对我国现有人工智能治理的政策和要求的落实与践行。例如，《倡议》在细则中特别强调了“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”。这个要求可以体现在我国刚刚发布的《科技伦理审查办法（试行）》之中。后者明确要求要对可能会带来科技伦理风险的科技活动进行审查监管，从审查主体、审查程序和监督管理分别提出了具体举措。

## 10. FTC 发布消费者对人工智能表示担忧的报告

发布日期：2023 年 10 月 3 日

来源：FTC 官网

链接：

- <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/10/consumers-are-voicing-concerns-about-ai>

摘要：

当地时间 10 月 3 日，美国联邦贸易委员会（FTC）发布了名为《消费者对人工智能表示担忧》的报告，重点关注了人工智能这一新兴技术对消费者和市场的风险问题，涵盖了数据、基础设施、数字系统的应用程序和设计等各个技术层面。

报告将消费者的担忧分为三类：第一，在如何人工智能构建问题上，FTC 表示训练人工智能模型所需的大量数据引发了消费者保护和竞争问题；第二，人工智能模型容易受到偏见、不准确及其糟糕表现的影响，因此消费者在使用人工智能模型时将主要担忧偏见和不准确，以及上诉途径有限这两个问题；第三，对于人工智能如何在现实世界应用的问题，消费者认为常见的网络安全投诉效果不佳，并提交了大量关于他们认为可能由人工智能驱动的诈骗和欺诈的投诉。FTC 表示，

未来他们将继续关注人工智能技术对于消费者的影响，他们的目标是防止随着人工智能技术的普遍化，消费者和市场可能面临更大的伤害。

### 植德短评

美国联邦贸易委员会（FTC）是美国联邦政府中负责消费者权益保护的企业监管机构，本份报告体现了 FTC 对人工智能的发展和消费者权益保护间关系的态度，即人工智能的发展不能影响到消费者权益。

FTC 此前的执法往往基于消费者投诉展开。本份报告总结了消费者重点关注人工智能对知识产权及生物识别数据、个人数据保护的影响问题，以及消费者目前面临的人工智能存在偏见和不准确性，及上诉途径有限这两个问题。FTC 也同样在报告中指出了人工智能技术应用可能带来的刑事法律风险，如网络钓鱼电子邮件、用于勒索的语音克隆诈骗等。

近日，我国国家金融监督管理总局也指出了 AI 新兴技术用于诈骗的刑事法律风险，就 FTC 指出的上述担忧，我国立法目前的规定是，要求保证训练数据来源的合法合规，畅通用户投诉渠道，注意防范生成式 AI 的虚假信息风险。相关服务提供者在注意防范生成式 AI 带来的刑事法律风险的同时，也需对上述合规要求及时自查，并按照相应法律法规履行合规义务。

## 11. 乌克兰发布《人工智能（AI）监管路线图》

发布日期：2023 年 10 月 7 日

来源：乌克兰数字化转型部官网

链接：

- <https://www.kmu.gov.ua/news/rehuliuвання-shtuchoho-intelektu-v-ukraini-mintsyfry-prezentuvalo-dorozhniu-kartu>

摘要：

当地时间 10 月 7 日，乌克兰数字化转型部在其官网公布了《人工智能 (AI) 监管路线图》（以下简称《路线图》）。该《路线图》旨在帮助当地企业为采用类似于欧盟人工智能法案的法律做好准备，并试图教育及引导企业更好地保护自己免受人工智能风险的影响

《路线图》设定了一个初步期限，使公司能够在未来两到三年内适应潜在的法律。同时，基于自下而上的方法，建议从少到多，并将为企业提供工具，以便在通过任何法律之前为未来的要求做好准备。

根据《路线图》，乌克兰人工智能立法草案预计将于 2024 年出台，不会在《欧盟人工智能法案》之前发布，以便使国家法规与地区法律保持一致。

### 植德短评

人工智能技术在全球的快速发展，引发了全球的广泛关注，使得各国政府争相、紧迫出台人工智能相关立法和政策，努力寻求创新的监管和治理方法，在积极利用人工智能和潜在的风险控制之间尽可能地实现平衡。

乌克兰发布《路线图》正是该趋势的重要体现。《路线图》体现出了乌克兰对于 AI 发展的积极态度，明确了人工智能对于乌克兰国家安全及未来发展的重要性，但也希望通过人工智能法律出台帮助企业 and 公民更好地防范人工智能风险。

同时，乌克兰也在《路线图》中指出，将努力实现未来国内人工智能立法与欧盟正在进行的《欧盟人工智能法案》无缝一体化衔接。《路线图》将为此奠定重要基础，如通过启动当地法律框架来指导当地公司的活动，以为欧盟立法做好准备。乌克兰的举措也是当前欧盟各国所作努力的重要体现。欧盟各国正在准备本国的人工智能法案，及相关监管举措，以为《欧盟人工智能法案》的正式出台扫清障碍、铺好道路。

## 12. 印度电子和信息技术部发布《印度人工智能 2023 计划》报告

发布日期：2023 年 10 月 14 日

来源：印度电子和信息技术部 (MeitY) 官网

链接：

- <https://www.meity.gov.in/content/indiaai-2023-expert-group-report-%E2%80%93-first-editionthe-ministry-electronics-and-information>
- <https://www.meity.gov.in/writereaddata/files/IndiaAI-Expert-Group-Report-First-Edition.pdf>

摘要：

印度电子和信息技术部 (MeitY) 于 2023 年 10 月 14 日发布了《印度人工智能 2023 计划》报告 (以下简称《计划》)。《计划》阐述了印度政府设想的印度人工智能战略的支点，以及对印度人工智能发展全面且野心勃勃的态度。

《计划》采用以任务为中心的方法，确保采取精确且有凝聚力的战略来弥合现有人工智能生态系统中的差距，详细介绍了建立卓越中心 (CoE) 进行日常运营及国家数据管理办公室 (NDMO) 负责管理数据收集、处理、存储等工作的制度工作，并重点关注了包括计算机基础设施建设、人工智能研究和创新能力提升、国家机器人战略草案草拟、人工智能芯片开发、印度数据集建设等问题，以最大程度地发挥人工智能的潜力，推动印度社会的进步。

### 植德短评

印度相较于欧盟和美国，在人工智能的监管上起步相对较晚。但印度政府在人工智能的发展上有着一颗野心，认为人工智能行业对于该国来说是一个“重要

和战略性”的领域。

此前，印度曾出台《国家人工智能战略》，提议为研究和采用人工智能开发一个生态系统，即#AI for ALL。印度电子和信息技术部也已经实施了“国家人工智能计划”。而《印度人工智能 2023 计划》提出了印度到 2026 年实现 10000 亿美元的数字经济目标，并明确了印度将制定国家机器人战略，相较于此前的战略和计划更为激进且涉及的具体领域更为广泛和细致，特别体现了印度政府希望全方位快速发展人工智能产业的态度。

同欧盟和美国相比，印度对待人工智能发展和安全平衡问题的天平更倾向前者，在监管层面更为宽松。本次《计划》也体现了这一点。《计划》基本忽视了风险的监管问题，专注于如何更好、更快地发展印度人工智能产业这一问题。未来，印度也将鼓励更多的人工智能企业进入印度市场，通过政府政策拉动人工智能产业的投资。

### 13. 美国、新加坡就人工智能发表联合声明

发布日期：2023 年 10 月 12 日

来源：新加坡通讯及新闻部

链接：

- <https://www.mci.gov.sg/media-centre/press-releases/singapore-and-the-us-to-deepen-cooperation-in-ai/>

摘要：

10 月 12 日，新加坡资讯通信媒体发展管理局（the Infocomm Media Development Authority of Singapore, IMDA）和美国国家标准与技术研究院（the US National Institute for Standards and Technology, NIST）在美国-新加坡关键新兴技术对话（the United States-Singapore Dialogue on Critical and Emerging Technologies, CET）中就两国人工智能治理框架联合行动发表联合声明。根据该声明，新加坡 IMDA 的人工智能验证（AI Verify）与美国 NIST 的人工智能风险管理框架（RMF）之间的联合映射工作已经顺利完成，从而实现了两国人工智能治理框架的互相可操作性。两国认为该行动是统一国际框架以促进可信和负责任的人工智能创新以及在国际上推进共同原则的重要一步。完成该项工作有助于为企业符合两国要求提供更加明确的指引，降低合规成本，并为人工智能部署和创新营造更有利的环境。

新加坡外交部还在本届 CET 上宣布成立双边人工智能治理小组，以推进人工智能方面形成共同原则和完成信息交流，并深化人工智能领域的研究和技术合作。

### 14. 英国人工智能安全峰会发布《布莱切利宣言》，包括中国在内的 28 国及欧盟共同签署

发布日期：2023 年 11 月 1 日

来源：英国政府

链接：

- <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

摘要：

2023 年 11 月 1 日，由英国主办的首届人工智能安全全球峰会在应该布莱切利园正式开幕。中国科技部副部长吴朝晖率领由科技部、外交部、中国科学院、阿里巴巴、腾讯组成的代表团参会，并在开幕式全体会议上发言。与会期间，代表团参与了人工智能安全等问题的讨论，宣介了中方几天前提出的《全球人工智能治理倡议》。

本次会议达成了关于人工智能国际治理的重要成果文件——《布莱切利宣言》（以下简称“《宣言》”），由包括中国、美国、英国在内的 28 国及欧盟共同签署。《宣言》称，人工智能(AI)带来了巨大的全球机遇：它有可能改变和促进人类福祉、和平与繁荣。人工智能应该以安全的方式设计、开发、部署和使用，是以人为本的、值得信赖和有责任的。人工智能带来的许多风险本质上是国际性的，因此最好通过国际合作来解决。《宣言》认为应对前沿人工智能风险的议程将侧重于识别共同关注的人工智能安全风险；以及确保安全的情况下，并酌情开展合作。

### 植德短评

《布莱切利宣言》是全球第一份针对人工智能发布的国际性声明。根据英国媒体评价，“这是罕见的全球团结表现”。虽然美国、欧盟、英国、中国等在人工智能监管的方针和优先事项上存在显著差异，且业界对 AI 治理方案争论激烈，但是包括上述国家在内的 28 个国家和欧盟一致认为，人工智能对人类构成了潜在的灾难性风险。中方也在会中积极宣介中方提出的《全球人工智能治理倡议》，再次明确了我国围绕人工智能发展、安全、治理三方面系统展开的人工智能治理中国方案，包括以人为本、建立风险评估体系、帮助发展中国家等在内的多条内容也被纳入了《宣言》之中。

《宣言》再次提示我们需要高度重视前沿人工智能及大模型的风险。研发端对确保人工智能安全具有特别重大的责任，因此前沿人工智能开发者在未来应以信赖和负责任的方式设计、开发人工智能，应用端也需避免人工智能滥用及虚假信息造成的风险。其中，尤其需要关注网络安全和生物技术等领域的风险。《宣言》要求各方通过法规等手段，防范和化解相关风险，公司也属于人工智能安全发展和治理的重要参与者，也可以在确保人工智能安全中发挥重要作用。



## 15. 环球音乐起诉人工智能公司 Anthropic 侵犯版权

发布日期：2023 年 10 月 26 日

来源：欧盟委员会官网

链接：

- [https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/universal-music-sues-ai-company-anthropic-copyright-infringement-levis-sues-coperni-trade-mark-2023-10-26\\_en](https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/universal-music-sues-ai-company-anthropic-copyright-infringement-levis-sues-coperni-trade-mark-2023-10-26_en)

摘要：

近期，包括环球音乐集团出版公司（UMGP）、协和音乐集团和 ABKCO 在内的音乐出版商已经对 Anthropic 提起诉讼。Anthropic 是一家由 OpenAI 前成员创立的人工智能公司，致力于开发大型语言模型。该诉讼于 10 月 18 日在纳什维尔联邦法院正式提起。

在诉讼中，环球音乐声称“在构建和运营人工智能模型的过程中，Anthropic 非法复制和分发了大量受版权保护的作品，包括出版商拥有或控制的众多音乐作品的歌词”。他们正在为每件侵权作品寻求 140,000 欧元的赔偿。他们还声称 Anthropic 的人工智能能够创建相同或几乎相同的标志性歌曲副本。根据出版商出示的起诉书内容，Anthropic 获得这些歌词的途径并不合法。根据 Anthropic 公司有限的披露，Claude 严重依赖的数据集（如 Common Crawl）包含来自大量网站的流行歌词内容，而这些歌词内容并没有获得出版商的许可。

### 植德短评

本案是 AIGC 普及以来的一次典型诉讼。双方的争议焦点主要在于 Anthropic 在构建和运营其人工智能模型过程中，是否非法复制和传播大量受版权保护的作品。本质上是训练数据的来源合法性及生成内容的版权保护两个问题。

在原告出版商们提供的长达 60 页的诉状中，他们明确提出，出版商拥抱“创新”以及道德和负责任地使用人工智能的潜力，但谴责违反版权法，使用受版权保护的原版歌词训练数据，并在生成内容中直接使用原版歌词，涉嫌侵犯原作者包括署名权、保护作品完整权、复制权、网络传播权等在内的多项版权权利内容。

在 AIGC 的浪潮下，可能引发 AI 产业发展和版权保护平衡的问题，并产生新型著作权侵权案件。该案的裁决结果也将为未来涉及人工智能的版权纠纷提供重要的法律指导和实践指引，帮助 AI 企业明确侵权的界限。未来，在音乐公司等版权方和 AI 公司间，也需要找到更好地合作模式，规避 AIGC 版权侵权风险的同时，实现作者、出版方、被许可人等之间的利益平衡。

## 特此声明

本刊物不代表本所正式法律意见，仅为研究、交流之用。非经北京植德律师事务所同意，本刊内容不应被用于研究、交流之外的其他目的。

如有任何建议、意见或具体问题，欢迎垂询 [aigc@meritsandtree.com](mailto:aigc@meritsandtree.com)。

北京植德律师事务所 元宇宙与数字经济委员会

AIGC 小组：时箫楠、王妍妍、赵芸芸、王艺

中国科学院大学经济与管理学院

孙毅教授课题组：孙毅、贺子涵、李子嘉、田章功

本期撰写人：王艺（植德）、田章功（国科大中丹学院）

## 北京植德律师事务所 元宇宙与数字经济委员会

## AIGC 小组合伙人成员介绍

**时萧楠**

合伙人/北京

**电话:** 010-5650 0937**手机:** 138 1006 8795**邮箱:** xiaonan.shi@meritsandtree.com**执业领域:** 知识产权、政府监管与合规、争议解决**工作经历:**

时萧楠律师是北京植德律师事务所合伙人。

时萧楠律师从事知识产权十余年，先后在北京天达共和律师事务所和日本西村朝日律师事务所、中国大型互联网公司工作多年，专注于解决合规、知识产权案件，包括互联网合规、数据合规、著作权授权、侵权诉讼、行政投诉等类型的案件，同时擅长解决疑难复杂案件。

时萧楠律师曾在大型知名互联网公司工作多年，对公司法务合规有着深刻的理解，并且深刻擅长以业务目标为核心提供解决方案。时萧楠律师有公司法务与律所双重经验，能以行业视角和律师视角多元提供知识产权纠纷、合规解决方案。

**代表业绩:**

- 知识产权：富士胶片专利许可相关合同纠纷（最高院商事法庭第一批案件）、易谱耐特软件著作权侵权、知名日本游戏公司与中国知名游戏公司著作权侵权
- 不正当竞争：站酷网
- 重大合规项目：知名APP合规评估；知名APP数据合规评估；各类型音乐曲库授权合作、投诉、维权应对；大型体育赛事合作；重大项目的著作权维权、维权应对；著作权集体管理组织合作等。

**教育背景:** 日本一桥大学，经营法（知识产权项目）硕士研究生



## 王妍妍

合伙人/北京

**电话:** 010-5650 0924

**手机:** 139 1089 6736

**邮箱:** [yanyan.wang@meritsandtree.com](mailto:yanyan.wang@meritsandtree.com)

**执业领域:** 投融资并购、银行与金融、政府监管与合规

### 工作经历:

王妍妍律师是北京植德律师事务所合伙人。在加入植德之前，王妍妍律师曾在北京市经纬律师事务所以及英国礼德律师事务所、美国杜威律师事务所等国际一流律师事务所工作数年，在投融资并购与跨境交易、银行与金融产品以及涉外争议解决等业务领域具有丰富经验。

王律师的主要执业领域包括投融资并购与跨境交易、银行与金融和争议解决，拥有丰富执业经验。曾代理过包括建筑、制造、新材料应用、银行、软件设计、文化娱乐、传媒、游戏、酒店、医疗设备、食品和体育等诸多行业的客户，对若干不同行业有深入了解，能根据行业特点为客户提供有针对性的优质法律服务，包括为这些客户提供融资，收购，公司治理、股权激励，架构重组等方面的法律服务。

### 代表业绩:

- 代表南山资本就投资镁佳科技、灵雀云、摩天轮、笑果文化、豹亮科技、不鸣科技、迷你玩、王牌互娱等TMT领域公司提供全方位法律服务
- 代表高榕、国开熔华产业投资基金完成对多个企业的投资
- 代表首旅置业处理其巴黎子公司参股酒店管理公司事宜以及参与境外基金投资及酒店改造项目提供法律服务
- 为中信银行参与的多项跨境银团贷款等事宜提供法律服务
- 为Terex Corporation、Nicklaus Company LLC（尼克劳斯）、Restaurant Brands International US Services LLC 等多家外资公司在中国的重组和经营提供法律服务

**教育背景:** 哥伦比亚大学，法学硕士

伦敦大学学院，法学硕士

中国政法大学，法学学士



## 赵芸芸

合伙人/北京

**电话:** 010-5650 0978

**手机:** 138 1160 9951

**邮箱:** yunyun.zhao@meritsandtree.com

**执业领域:** 投融资并购、争议解决、政府监管与合规

### 工作经历:

赵芸芸律师是北京植德律师事务所合伙人，曾先后在北京大成律师事务所律师、北京市天银律师事务所执业，并自2010年起任北京华录百纳影视股份有限公司证券部负责人、法务部总监，同时兼任海外事业部、丹丹百纳经纪公司及运营协调部负责人，后加入北京植德律师事务所。

赵芸芸律师早期从事 IPO/MA 等资本市场领域的非诉业务及股权纠纷类仲裁业务，后致力于文化娱乐、影视传媒及 TMT 领域。赵芸芸律师在国内及中外合作电影（含动画电影）、电视剧、综艺栏目、体育赛事、杂志媒体合作、网络游戏运营等项目类方面，为客户全程提供商业合作模式架构设计、谈判策略与要点、合规风险评估与解决等法律服务，并在融资安排、联合投资合作、发行、商务模式等有独特心得与经验，善于根据不同项目有针对性地提示风险点和设计协议。在资本市场领域，赵芸芸律师曾多次作为被投资方专项法律顾问，与投资方就公司估值、著作权尽职调查、业务增长模式、有限合伙等持股平台设计、员工激励与不竞争、SPA 协议、SHA 协议等进行沟通、谈判并形成一揽子协议，有效维护客户最大估值利益及股权结构安全性。

### 代表业绩:

- 文化娱乐及消费：参与或经办《建国大业》《深夜食堂》等多部电影项目、《媳妇的美好时代》《双面胶》等多部电视剧项目、《笑傲江湖》等多部电视栏目并提供法律服务，代表客户获得《T》《Nylon》《Wallpaper》等杂志授权发行；
- 资本市场：华录百纳创业板 IPO 上市、胖虎奢侈品上市前红筹重组及 C 轮融资、代表深蓝影业及栩栩华生接受华人文化投资、代表数字栩生接受蓝色光标投资等，并代表华录百纳、栩栩华生、磨铁等企业客户经办其投融资并购业务及证券相关业务。

### 荣誉奖项:

赵芸芸律师曾上榜 2019 年度、2020 年度、2021 年度 LEGALBAND 中国顶级律师排行榜“体育娱乐”版块。

**教育背景:** 中国政法大学，法学硕士



## 王艺

合伙人/深圳

**电话:** 0755-3325 7513

**手机:** 136 3157 5683

**邮箱:** yi.wang@meritsandtree.com

**执业领域:** 政府监管与合规、投融资并购、争议解决

### 工作经历:

王艺律师是北京植德（深圳）律师事务所合伙人、植德公司合规部负责人。在加入植德之前，王艺律师曾在广东华商律师事务所工作9年并担任高级合伙人。

王艺律师主要执业领域为政府监管与合规（数据合规为主）、投融资并购与争议解决，拥有十多年的丰富执业经验，曾代表多家境内外上市公司、知名企业提供数据合规、金融科技法律服务，行业领域包括金融、汽车、医疗、云服务商、科技、酒店、制造业、设计等行业，为众多高成长性公司提供数据合规投资端、资产端、交易端、争议解决端的法律服务。此外，王律师还为深圳地方金融监督管理局、前海管理局、深圳市工业设计行业协会、杭州全球金融中心等政府部门、行业协会等提供法律服务。

### 社会职务:

- 武汉仲裁委员会仲裁员
- 深圳市工业设计行业协会监事会主席
- 信通院“数据安全推进计划（DSI）”数据安全专家、中国信通院个人信息保护合规审计推进小组专家成员之一、信通院卓信大数据计划数据安全培训讲师
- 中国互联网金融协会统计分析专业委员会委员
- 全国金融标准化技术委员会秘书处专家组成员

**教育背景:** 香港大学，数字化转型与企业战略 研究生  
深圳大学，经济法 法学硕士、法学本科

## 中国科学院大学经济与管理学院

## 孙毅教授课题组成员介绍



## 孙毅

中国科学院大学经管学院教授、博导

邮箱: suny@ucas.ac.cn

**研究领域:** 数字经济、数字化转型、数据要素市场

**研究经历:**

孙毅, 中国科学院大学经济与管理学院教授、博导, 院长助理; 数字经济与虚拟商务系副主任; 中国科学院大学继续教育学院副院长、培训中心副主任; 第二届全国基层政权和社区建设专家委员会青年委员; 新华社特约经济分析师; 阿里研究院活水学者理事会理事; 神思电子(300479)、尚航科技(836366)独立董事。

主要研究领域为数字经济、数字化转型、数据要素市场、金融科技、智能制造。出版《数字经济学》教材; 先后主持科研30余项各类课题, 其中包括国家自然科学基金3项, 中科院各类重点课题5项, 国家发改委、财政部、科技部、教育部、中国科协等中央部委课题6项, 北京市、山西省经济发展重点规划课题2项, 百度、阿里巴巴、广发证券等大型企业合作课题3项, 在国内外重要期刊发表论文30余篇。多次参与国家数字经济领域政策的修订和研讨, 并参与多个省市数字经济发展规划。课题成果“数字经济监测指标体系”被北京市委政府确定为“十四五”北京市全球数字经济标杆城市建设工作的评估标准。主持的多项课题成果已在百度、广发证券等企业得到推广和应用。先后接受新华社、人民日报、央视财经频道、北京卫视财经频道等主流媒体的约稿、访谈, 发表关于数字经济、人工智能和大数据产业发展的相关观点。

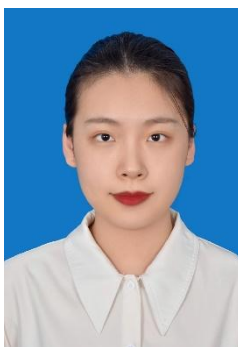


## 贺子涵

中国科学院大学经管学院硕士研究生

邮箱: [hezihannihao@163.com](mailto:hezihannihao@163.com)

研究方向: 数字经济、数据要素市场化机制设计



## 李子嘉

中国科学院大学中丹学院硕士研究生

邮箱: [qursin@outlook.com](mailto:qursin@outlook.com)

研究方向: 数字经济、企业数字化转型



## 田章功

中国科学院大学经管学院博士研究生

邮箱: [tianhao94@163.com](mailto:tianhao94@163.com)

研究方向: 数字经济、数字化就业





人工智能月报系列 请扫码阅读



图灵财经公众号



植德公众号