相关法规亟待“靴子落地”

面对当前 AI 行业的发展乱象，安全治理迫在眉睫。

2022 年 3 月开始实施的《互联网信息服务算法推荐管理规定》，就把生成合成类算法作为五类算法推荐技术之一纳入监管。

2022 年 12 月施行的《反电信网络诈骗法》明确了互联网企业的风险防控责任，平台有义务针对 AI 诈骗进行技术筛查、拦截和报告。

今年 1 月起施行的《互联网信息服务深度合成管理规定》中明确了：任何组织和个人不得利用深度合成服务制作、复制、发布、传播法律、行政法规禁止的信息，并要求深度合成服务提供者应当建立健全辟谣机制。这为“AI 换脸”技术应用划定了“底线”和“红线”。

但我国目前仍然缺乏对人工智能的总体立法，以及人工智能伦理的行业规范。

4 月 11 日，国家互联网信息办公室发布了《生成式人工智能服务管理办法（征求意见稿）》（下称“征求意见稿”），并公开征求意见。征求意见稿对隐私、知识产权、肖像权、名誉权、训练数据、人工标注、标识、不公平竞争、防沉迷等提出了规范。

这是国家首次针对当下爆火的生成式 AI 产业发布规范性政策。具体来看，征求意见稿制度设计，一方面延续了《互联网信息服务深度合成管理规定》关于标识、评估、备案等既有制度安排，另一方面就训练数据、人工标注等提出了新的合规要求，体现出监管逻辑更新和制度体系迭代的新趋势。

上海正策律师事务所董毅智律师告诉《新民周刊》记者，这是对人工智能进行管理的良好开端。而《人工智能法》也已列入立法计划，草案预备年内提请全国人大常委会审议。近日，《国务院 2023 年度立法工作计划》对外公布，披露了这一动向。这表明，中国官方将推动全国层面的人工智能专门立法。但真正落地还需要一段时间。在未来，仍然需要制定更详细的规定，对使用人工智能的各个主体进行更加明确的责任认定，对提供技术的服务商、内容生产者和平台等进行责任区分。

“从监管角度来说，首先要明确监管机构，是不是两会后新组建的国家数据局？其次，要加强行业协会的自律，提供深度合成服务的企业要严格落实信息安全管理主体责任，建立健全算法机制审核、科技伦理审查、反电信网络诈骗等管理制度，加强合成内容的审查与管理，对智能对话、合成人声、人脸生成、沉浸式拟真场景等生成内容进行显著标识，避免公众混淆或误认。”董毅智表示，对于用户来说，要提高使用门槛，“至少必须实名制”。

此外，还要压实网络平台的责任。对违规的发布者、平台予以惩戒；作为个人信息的重要处理者，承担起保护个人信息的法定义务。就在 5 月 9 日，抖音发布了一则针对人工智能生成内容的倡议，禁止利用生成式人工智能技术创作、发布侵权内容，包括但不限于肖像权、知识产权等，一经发现将严格处罚，此外还要求发布者对人工智能生成内容进行显著标识。

“当然我们每个人都要提高安全意识，保护好个人隐私，不要在公开平台上随意发布自己的清晰正脸照，尤其是未成年人的。”董毅智强调。

新兴领域立法速度加快，已经成为了世界各国的一个常态。针对 AIGC 带来的安全风险，董毅智表示，美国已经约谈了几家巨头，但总体偏向于鼓励；欧洲的监管则相对较严，正在加紧出台相关法律法规，“我国差不多介于两者之间。这项技术因为太新了，目前也尚未出现判例。整个行业可能还要经历一段时间的野蛮生长”。

当前“治理”依然赶不上“诈骗”技术更迭的速度。长期而言，就需要采取个人预防、防御技术迭代、压实平台责任、完善法律法规等多管齐下的方式。但不可否认的是，掌握人工智能治理的话语权、规则制定权，抢先形成新的国家竞争优势，已经成为世界各国努力的目标。

左图：许多基于人工智能技术的应用程序也层出不穷，甚至包括变脸软件和“一键脱衣”等应用程序。