

# 算法无处不在 如何预防风险

进入大数据时代以来，算法几乎无处不在。从银行贷款到汽车保险，从医疗信息到社交数据，从智能写作机器人到ChatGPT，算法也越来越深入而广泛地引导和影响着我们的生活。各国都在探索对于算法的有效监管路径，以促进相关产业的繁荣发展。在世界范围内，欧盟与美国是较早对算法进行专门立法规制的国家和地区，分别代表了两种不同的治理机制。本期“域外之音”对这两种算法治理机制进行了简单的介绍。



## 个体赋权与外部问责的分立

### 欧盟：

#### 数据治理框架下的个体赋权机制

2018年5月25日生效的《欧盟数据保护通用条例》构建了欧盟的数据治理框架和核心制度，通过赋予数据主体新型权利的方式将其算法治理模式嵌套其中，以数据治理与算法治理联结的方式初步构建了针对用户画像和自动化决策的规制机制。

《欧盟数据保护通用条例》赋予数据主体对数据处理和数据画像等行为的反对权、知情权、访问权、更正权和删除权等，要求企业在利用用户行为信息等进行进一步处理和个性化推荐、自动化决策等行为时，应当告知用户对数据处理的事实、原因、基本算法逻辑、预期分析后果和可能产生的风险，且用户有权免受完全基于算法对数据的自动化处理（包括用户画像）作出的具有法律效力或可能产生重大影响决定的制约。

与此同时，《欧盟数据保护通用条例》在引言中，对于用户免受完全基于算法自动化处理决策影响的权利作出进一步的细化，规定了数据主体享有对自动化决策获得合理解释以及对相关决定进行质疑的权利。可以说，《欧盟数据保护通用条例》希望通过赋予个体对数据的控制权，进而达到对算法应用施加干预的效果。

### 美国：

#### 公共治理框架下的外部问责机制

由于法律传统和核心理念上的差异，欧盟和美国构建了不同的算法规制机制，也成为算法规制领域两种颇有代表性的治理体系。

欧盟将个人自治权与人格尊严的保护作为核心理念，将个体赋权与算法流程控制相结合，构建了数据治理框架下的个体赋权模式；而美国将反算法歧视和决策公平性作为核心理念，将外部问责和行政监管相结合，构建了公共治理框架下的算法问责模式。

美国率先在公共领域通过组建专门机构和人员构成问责主体的方式建立了外部问责机制，并通过一系列法案完善问责机制，确保其有效运行。

2017年，美国《关于算法透明性和可问责性的声明》就明确使用算法决策的机构应当对算法所遵循的程序和具体制定的决策进行解释，同时在包括系统开发在内的每个阶段应用算法透明度和可问责性的原则，以及对数据、算法模型和决策等过程进行记录，以便于后续产生损害时的算法审计。

2019年，美国《算法问责法》将消费者保护和算法规制作为立法宗旨，要求大型算法平台（主要是指拥有或控制个人信息较多的平台主体）对自动化决策给消费者带来的信息隐私和安全风险以及因种族、宗教、政治信仰或性别等方面的差异而造成的算法歧视与偏见进行评估，并采取消除此类风险和歧视。该法案将自动化决策定义为“通过一个包括机器学习、数据分析、处理以及人工智能技术的计算机程序作出的与特定消费者相关的决策”。

2021年，美国又通过行政命令的方式，要求司法部和联邦贸易委员会采取措施，限制不良垄断行为并制定针对大型互联网企业的数据监管规定。

总之，美国算法问责机制的核心在于算法平台有义务对其算法系统设计和决策应用的合理性进行报告和证明，并有义务减轻算法可能带来的负面影响和潜在风险。

## 个体赋权与外部问责的兼容

值得注意的是，近年来两种制度出现了兼容以至相互吸纳的现象，也有一些共性发展的趋势和做法。

### 欧盟：以算法问责辅助个体赋权

2019年4月，欧盟《算法问责及透明度监管框架》就算法自动化决策应用提出了全面的监管要求，将算法透明度与可解释性作为实现算法公平的基础性条件，并就“建立适当的伦理和法律框架”明确了具体的规制路径和政策建议。

2020年12月，欧盟《数字市场法》和《数字服务法》制定了更加适应平台和数字市场的竞争监管规则，进一步明确数字服务提供者的责任、遏制平台恶性竞争行为，基于大型平台企业具有制定市场准入门槛的能力，对这类“市场守门人”提出“守门人法规”，以缓解数字经济时代大型平台与用户之间的权利失衡，实现用户权利保护和公平市场的形成，明确算法问责和审计规则，实现算法透明度的提升。

2022年11月，欧盟《数字服务法》对算法透明度进行了更深入的探讨，采取算法问责和透明度审计等措施规范平台算法使用和推广行为，要求在线平台公开推荐活动所依据的算法参数等以提升算法透明度，并成立欧洲算法透明度中心，针对算法评估提供新的透明度评估方法，以探索最佳实践路径。

此外，《2020年新西兰算法章程》也通过外部问责方式提升算法透明度，要求签署该章程的政府机构必须保证算法驱动决策的过程公开、透明、道德，包括提供算法过程和存储相关信息、进行风险评级以评估偏见出现的可能性及影响程度等，以构建公众对于政府机构自动化决策行为的信任。

### 美国：以个体赋权强化外部问责

2020年6月，《算法正义和互联网平台透明度法案》进一步明确算法自动化决策的监管规则，赋予消费者对数据收集的异议权和要求对算法自动化决策进行审查和解释的权利，提出披露算法收集个人信息的类别以及处理个人信息的方式、发表透明度报告等方式确保平台算法透明义务得到履行。

2022年，《算法责任法案》再次重申将提升算法透明度作为重点，要求算法平台对算法可能造成的偏见进行评估，以消除对少数群体存在的歧视。各州政府立法也通过要求披露等方式提升算法透明度，例如2022年华盛顿《停止算法歧视法案》要求平台披露使用算法的情况，并在作出不利于消费者的决定（例如拒绝抵押贷款等）时，消费者有权获得更强大的解释说明；佛罗里达州众议院《技术透明度法案》也明确消费者有权要求披露、删除或更正数据，并选择加入或退出数据的销售或共享，未经消费者授权，数据控制者不得通过语音识别功能收集消费者的精确地理位置数据或个人信息等。

### 从事后问责到风险预防的转向

2021年4月欧盟《人工智能法案（草案）》以风险评估和划分的视角，对平台企业和政府如何合理规范使用算法进行了规定，采取分级分类的思路，以算法可能对基本权利造成的损害程度为依据，将算法的使用划分为不可接受的风险、较高风险、有限的风险和极小风险四类，并对高风险任务设置了包含第三方评估、建立风险管理系统等在内的监管措施，设定较高的罚则，规定全面、范围广泛。该法案将风险等级分类作为基础，以明确人工智能技术可能对个体健康、安全或基本权利造成的风险水平，同时设置了严厉的违规处罚措施，进而实现人工智能发展与安全的平衡。

2022年9月，欧盟《AI责任指令》引入“因果关联推定”和“获取证据的权利”两项要素，以帮助AI系统的受害者证明他们的责任主张。前者是指当平台的过错与人工智能的效能之间产生因果关联具有合理性时，受害者在证明平台过错之后，无需再对过错与损害之间的因果关系进行详细证明。后者是指在高风险算法系统具有引起损害的可能性时，司法机关可以在原告请求下，指令算法系统的开发者和使用者披露相关证据。这两项规则能够减轻“算法黑箱”造成的不透明性和不可预见性，减轻原告举证责任，确保由于人工智能造成损害的受害者能够与其他技术造成损害的受害者获得同样有效的赔偿，以实现权利救济和风险救济的有效性。

作为与现实生活深度融合的智能技术，算法对现代社会秩序造成的冲击不可忽视，确保算法透明、公平、负责，对人工智能技术健康发展具有重要意义。随着算法治理的不断深入，单纯的个体赋权机制或者算法问责机制都难以应对技术飞速发展带来的冲击，两种机制开始出现融合，并促使算法透明度提升以及风险预防成为治理重点。

（来源：人民法院报）

