美参院将评估 AI 算法偏见提上立法日程：“算法黑箱”迎来光明？

人工智能1 周前

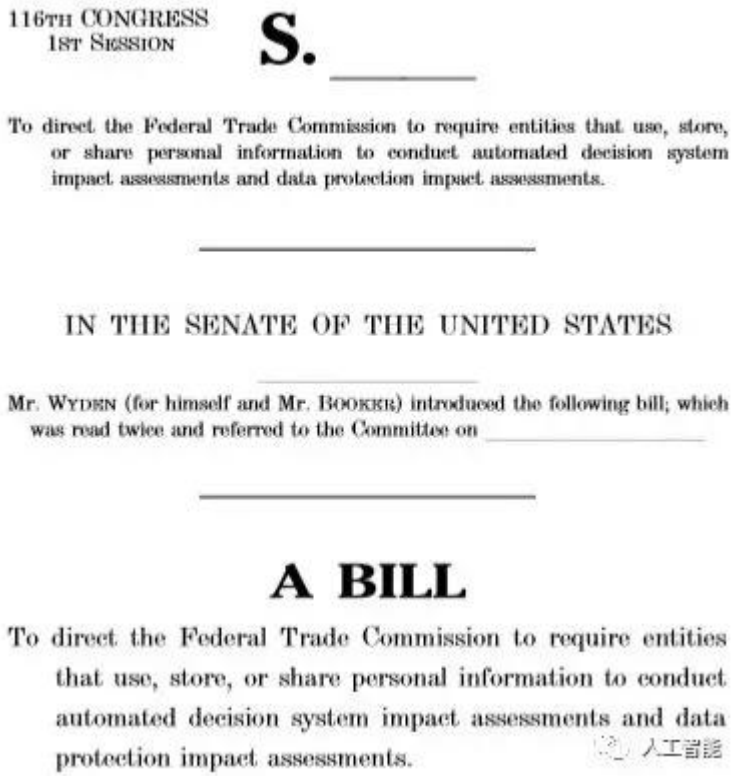
“当社会达到万物互联的普适计算时代后，一切都被数据化，所有的社会运行可能都是通过算法（自动决策系统）在支配。到那时算法（自动决策系统）将成为社会基础设施的一部分。所以算法阳光化是势在必行的，透明化是通往可理解的必要路径。”

4 月 1 日 ,美国 2020 年总统竞选候选人 ,参议院布克在华盛顿参加“We The People”集会。

人工智能机器学习所做的自动化决策是否客观？如果算法带有偏见，在使用人脸特征等个人信息后，科技公司或其他实体需要对其算法产生的决策结果负责吗？

当地时间 4 月 10 日，美国两位民主党参议员布克（Cory Booker）和怀登（Ron Wyden）联合提出了《2019 算法问责制法案》（Algorithmic Accountability Act of 2019），试图对人工智能机器学习中的偏见和个人敏感信息使用问题进行规制。

法案提出，联邦贸易委员会 (Federal Trade Commission) 应尽快制定关于“高风险自动决策系统”（high-risk automated decision system）的评估规则，科技企业必须评估算法是否存在歧视性偏见，以及它们是否对消费者构成隐私或安全风险。



116TH CONGRESS
1ST SESSION

S. _____

To direct the Federal Trade Commission to require entities that use, store, or share personal information to conduct automated decision system impact assessments and data protection impact assessments.

IN THE SENATE OF THE UNITED STATES

Mr. WYDEN (for himself and Mr. BOOKER) introduced the following bill; which was read twice and referred to the Committee on _____

A BILL

To direct the Federal Trade Commission to require entities that use, store, or share personal information to conduct automated decision system impact assessments and data protection impact assessments.

人工智能

《2019 算法问责制法案》第一页截图。图片来源：美国国会

据科技媒体 The Verge 报道，此项法案主要针对的是在数据领域获得大量个人信息的科技巨头们。年收入超过 5000 万美元，拥有至少 100 万个人或设备信息的公司，或主要业务就是买卖个人数据的科技公司将属于这项法案主要的规制目标。

2020 年美国总统参选人布克在声明中说，必须采取更有力的措施解决“恶性”歧视在科技平台上的使用，虽然有时候这些科技平台是无意的。

法案的另一发起人怀登对美联社表示：“计算机正在越来越多地参与到美国人日常生活中的关键决策。例如申请人是否可以买房、找什么工作，甚至是否应将某人送进监狱。当这些公司（根据法案）开始评估和调查的时候，他们将在系统中发现很多带有偏见的决策意见。”对于人工智能公司在透明度和公开性方面做出的努力，怀登表示：“很明显，自我监管在这里已经失败。”

美国参议员怀登（Ron Wyden）。

在这项法案中，“高风险自动决策系统”主要指涉及种族、肤色、民族、政治见解、宗教、工会成员、遗传数据、生物特征数据、健康、性别认同、性取向、刑事定罪或逮捕的个人信息的系统。当这些系统试图分析或预测个人的工作业绩、经济状况、健康、个人偏好、利益、行为、所在地点或行动，从而改变个人的合法权利的时候，或当这些系统被用于监测一个大的公共场所时，它们必须被相关法案评估和监测。

对此，科技公司支持的智库“信息技术和创新基金会”(Information Technology And Innovation Foundation) 副主卡斯特罗 (Daniel Castro) 对法案表示了不满。他说：“如果某个决定有很高风险损害消费

者利益，那么无论是算法还是个人做出这个决定都是一样的。用更高标准要求算法意味着自动化决策在本质上比人类决策更不可信或更危险，但事实并非如此。”

卡斯特罗还表示，这项法案对大公司很不公平。如果公司需要对每一次软件更新都进行评估的话一定会损害产品开发过程。

美联社认为，这项法案的提交表明，对数据经济加强管理正逐渐成为受到两党支持的重大议题。一旦法案在国会通过，从社交媒体、在线数据经纪公司、金融算法到自动驾驶软件等都将受到影响。

另外，在“高风险自动决策系统”使用的所有信息中，生物特征是敏感度和风险系数最高的信息之一。在商业领域使用日渐广泛的人脸识别技术更是让美国的立法者产生了很大担忧。就在上个月，民主党参议员施茨（Brian Schatz）和共和党参议员布朗特（Roy Blunt）就提出了《2019 商用人脸识别隐私法案》（Commercial Facial Recognition Privacy Act of 2019）。根据这项法案，公司在公共场所使用面部识别技术之前，以及在与第三方共享任何数据之前，都必须征得人们的同意。

它还将要求评审机构在产品使用之前对其进行测试，努力在前消除算法偏差和准确性方面的问题，以及要求面部识别技术服务商达到联邦贸易委员会和国家标准与技术研究所制定的安全标准。

那么，算法偏见在实际操作中又是如何产生的呢？

一方面，机器学习算法相关细节一般属于专利保护范围，因此被科技公司严格保密。另一方面，算法本身可以复杂到连代码创建者都不知道它们是如何工作的。这就是所谓的“黑箱”问题。

科技杂志《未来主义》的作者戈利普尔（Bahar Gholipour）曾在去年撰文指出，对于算法“黑箱”的担忧在计算机科学家群体中已经流传多年，而在过去的两年中，关于人工智能公平性的论文数量激增，反应了人们对这个议题日渐增长的兴趣和关心。

《未来主义》戈利普尔的长文页面截图，图片来源：未来主义

湖南师范大学人工智能道德决策研究所博士，上海玛娜数据科技发展基金会研究员胡晓萌对界面新闻表示：“（算法）技术人员通常在自动决策系统的设计、研发、运营等过程中负责一个子集，很少有人可以了解技术的全貌。没有人了解，也就意味着无人能牢固地、抑或有效地控制技术（的使用和发展）。”

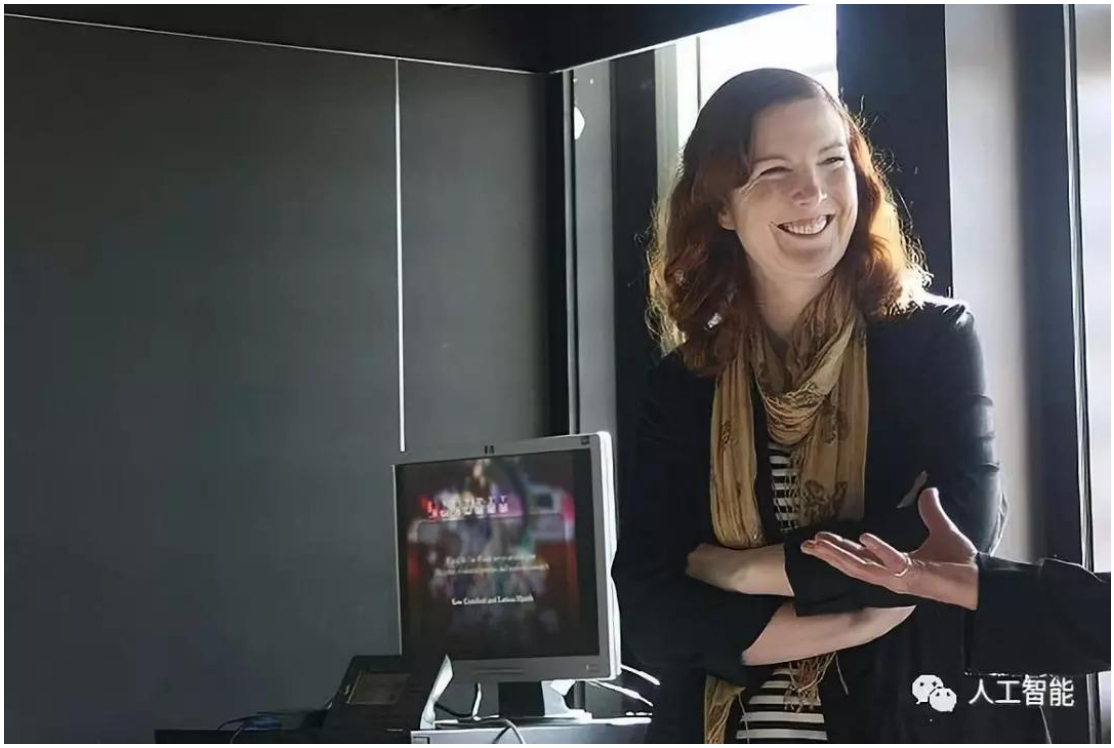
而由于机器学习算法又是基于历史数据之上的产物，胡晓萌认为“算法会不可避免地将价值观和伦理带入产品和服务的设计过程中，将价值观蕴藏于商业决策之中。对于企业来说，考量价值观和伦理的目的不仅在于解决当前问题和带来丰厚收益等，更在于决策本身是否会被企业自身和用户接受、算法输出的结果是否会被用户信赖。”

而对于算法偏见对社会的整体影响，戈利普尔在文章中称：“如果这一问题不能得到解决，那它将会摧毁我们的社会。因为这些算法将确保我们努力挣脱的所有歧视被永久地编码到我们的未来之中。”

科技杂志《连线》也曾在 2017 年发表文章表示，机器学习的自动决策机制对美国宪法所保障的“正当程序权”（due process）产生了威胁。

文章援引纽约大学著名人工智能研究机构 AI Now 的报告说，负责刑事司法、卫生和福利等领域的公共机构正越来越多地使用评分系统和软件来指导或决策重大事件，如准予保释、判刑、执法等。但政府在使用这些算法时应当慎重。

AI Now 的创始人克劳福德（Kate Crawford）说：“在算法做出的自动化决定中，我们应当为人们提供和人类决定一样的正当程序保护。”



纽约大学人工智能研究机构 AI Now 联合创始人克劳福德

布克和怀登在 4 月 10 日提出的法案并非立法者们第一次尝试对算法“黑箱”问题进行规制。2018 年 5 月 25 日，《欧盟一般数据保护条例》（GDPR）生效，其中最具争议性的条款之一就是消费者有自动化决策“解释权”，并且在某些条件下，消费者有权利不接受完全由（人工智能）自动化系统做出的重大决定。在这项法案生效前，全球消费者没有任何法律工具对人工智能算法作出的自动化决策进行挑战或提出异议。

那么，算法“黑箱”的不可解释性是否在技术上完全无解呢？

杜克大学计算机科学、电气和计算机工程副教授辛鲁丁(Cynthia Rudin) 曾对《未来主义》 表示，计算机科学家们拥有建立非“黑箱”的算法模型的能力。他说：“但是，要让人们注意到这项工作有一定困难。如果政府机构不再为‘黑箱’算法买单，那将会有所帮助。如果法官拒绝使用‘黑箱’算法进行判决，这也会有所帮助。”

对此，胡晓萌认为，在算法时代，社会的权力运行就在黑箱之中，因此，“对算法进行法治的目的就在于让社会权力运行阳光化”。

他表示：“当社会达到万物互联的普适计算时代后，一切都被数据化，所有的社会运行可能都是通过算法（自动决策系统）在支配。到那时算法（自动决策系统）将成为社会基础设施的一部分。所以算法阳光化是势在必行的，透明化是通往可理解的必要路径。