

欧盟环境法之预防原则在自动化决策当中的应用及对我国的启示

袁宇

贵州大学法学院, 贵州 贵阳

收稿日期: 2023年3月13日; 录用日期: 2023年4月7日; 发布日期: 2023年5月31日

摘要

随着自动化决策技术的发展, 其运行模式更加接近于由机器在数字环境中自行作出决策行为, 并对人类生活产生直接影响, 且此种影响难以被预见。该模式类似于人类对自然环境产生了不可预知的损害。基于损害程度的不可预知性, 可将其与欧盟环境法中的预防原则的相似性进行比较分析。欧盟的《通用数据保护条例》(General Data Protection Regulation, 以下简称GDPR)中, 已然将欧盟环境法中某些措施予以吸收借鉴, 例如数据保护的影响性评估手段(Data Protection Impact Assessment, 以下简称DPIA)。基于分析预防原则对欧盟数字立法的影响, 对促进我国《个人信息保护法》的修改和完善具有积极影响。

关键词

自动化决策, 数字环境, 预防原则

The Application of the Precautionary Principle of EU Environmental Law in Automated Decision-Making and Its Implications for China

Yu Yuan

School of Law, Guizhou University, Guiyang Guizhou

Received: Mar. 13th, 2023; accepted: Apr. 7th, 2023; published: May 31st, 2023

Abstract

With the development of automated decision-making technology, its operating mode is closer to-

machines making decisions on their own in a digital environment and having a direct impact on human life, and this impact is difficult to predict. This mode is similar to humans causing unpredictable damage to the natural environment. Due to the unpredictability of the degree of damage, it can be compared and analyzed with the similarity of the precautionary principle in EU environmental law. In the EU's General Data Protection Regulation (GDPR), some measures in EU environmental law have been absorbed and borrowed, such as the Data Protection Impact Assessment (DPIA). Based on the analysis of the impact of the precautionary principle on EU digital legislation, it has a positive impact on promoting the revision and improvement of China's Personal Information Protection Law.

Keywords

Automated Decision-Making, Digital Environment, Precautionary Principle

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

自动化决策是与自然人决策相对立的概念，意指利用计算机技术、算法程序、深度学习或神经网络替代自然人处理关键数据，凭借其自动生成对数据主体具有法律效果的决策的行为[1]。自动化决策在当今世界对民众的影响越来越大，它们如今被广泛应用在司法行政或公共服务当中。然而自动化决策的高效便民往往是建立在缺乏透明度的前提之下，民众不仅无法预知自动化决策会将其生活导向何方，对自动化决策将会产生的社会整体影响更加缺乏可预知性。欧盟在数字领域的立法是较为先进的，2016年的《通用数据保护条例》(General Data Protection Regulation, 以下简称为 GDPR)已提出相应的监管措施如数据保护影响性评估(Data Protection Impact Assessment, 以下简称为 DPIA)，以应对实施自动化决策所带来的风险挑战。DPIA 的提出，借鉴了欧盟环境法中的预防原则，即在缺乏足够的科学技术知识的情况下，不允许以此为借口延误时机，应在费用可接受的范围内，对可能给环境造成重大损失的、无法避免地可预见灾害及时采取适当的防治措施。该原则本适用于环境保护、消费者保护和健康保护等领域。如今可将其扩大解释到数字领域的风险之中，即给提供自动化决策服务的主体提出了先行评估、自动防治的义务。

本文通过分析欧盟环境法中的预防原则中，面对未知的(数字)环境风险挑战，和缺乏可预测的科学证明资料前提下，比较人类干预对自然环境造成的危害与自动化决策技术的发展对人类生活的某些方面造成的危害之间存在的相似之处，即在面对缺乏科学的确定性和可预测性前提下，如何通过事前立法化解风险挑战，从而对我国《个人信息保护法》的修改与完善提出合理建议。

2. GDPR 中的自动化决策规制方式及其不确定性分析

2.1. 自动化决策概念

自动化决策一词于《通用数据保护条例》(GDPR)中第 22 条被提及，《GDPR 中文版》第 22 条第 1 款规定：“数据主体有权利不受一个仅仅依靠包括分析的自动化处理的决定的限制，这会产生关于他/她或仅仅影响他/她的法律或类似后果[2]。”通过 GDPR 第 22 条的规定，欧盟定义了使用自动化决策的可能性边界。它对仅以机器算法作出的对个人产生重大法律后果的行为予以禁止，但从法理学视角俯瞰这一规定，能发现该规定的术语解释存在很大的讨论空间。条文规定不明晰，将会在实际规制方面产生困

境，不利于自动化决策本身的解释及发展。

2.2. 自动化决策规制的不确定性

首先，条文中对“仅仅”(Solely)、“决定”(Decision)、“法律或类似后果”(Legal or Similarly Significant Effect)等法律术语的解释尚待清晰，其次，GDPR 第 22 条第二款对数据权利主体的拒绝权进行了限缩，该款的作用即为“但书”。也就是在数据主体同意、为双方合同履行之必要、欧盟及其成员国法律强制规定的事项出现时，数据主体无权行使其拒绝权。“但书”的出现，在民众心中对 GDPR 能否切实解决好自动化决策带来的人格权保护挑战又增加了疑问。

此外，根据 GDPR 第 13 条第二款、第 14 条第二款和第 15 条第一款规定，在民众信息自动决策的情况下，数据控制者有义务向数据主体提供控制者意图将个人数据向第三国或者国家组织进行传输的事实、委员会是否就此问题做出过充分决议、第 46、47 条或者第 49 条第 1 款第二小段提及情形的相关信息、数据控制者的身份和联系方式、数据收集的方式和目的等从数据收集，数据处理以及数据流向一系列具有内在逻辑的，庞杂的信息。相关文献将信息主体的这一权利称为“解释权”[3]。该规定要求数据处理者将这一些列数据流以数据目录的形式提供给信息主体，其中既包括来源于该主体的个人信息，也包括从其他主体收集但可以辅助对该主体的信息进行处理的其他信息，这也是主体知情权的体现。但是 GDPR 所规定“解释权”本身便是一个受到广泛学术讨论的主题，既有学者发表支持观点[4]，也有反对的声音[5]。矛盾源于对这些条款解释存在的一些疑问，例如：数据控制者应当向数据主体公开决策的逻辑链条中哪个部分的信息？如果算法黑箱中的决策逻辑连数据提供者也不能清晰认知，那么应当如何保证数据主体的知情权？这些问题的根源在于人工智能和机器深度学习的飞跃式发展，人们普遍质疑使用这些技术做出决策的可预测性。同时，实现自动化决策的目标之一便是授予算法一定程度的自主权。决策过程自动化的目的是使算法获得比人类参与者执行的相同活动更高效的结果。为了做到这一点，有必要允许编程工具分析数据，并在最先进的人工智能技术中自主学习。

欧盟各成员国法院和监管机构从人权角度对 GDPR 自动决策条款进行了解释，诸如对“法律效力”等要素的解释从人权的角度来看，可以指对公民健康权的任何限制[6]。或者“有价值的信息”一词，应采用对文义最广义的理解，扩大条款的适用范围，从而规避条款的风险[7]。即使这些解释能够在一定范围内解决 GDPR 条款由于解释不清而难以适用的问题，但仍未给予该条款列举式的规定，使得条款的适用仍然归于抽象，并未解决其核心的潜在危险性问题。首先，自动化决策的结果仍然难以预测，并包含潜在危害性；其次，对这些规定的解释仅仅针对公民个体的权利，未能考虑到对社会或国家具有重大影响的数据决策成果，即忽略对公共安全的考量；其三，GDPR 条款存在固有性缺陷，其并未作兜底规定，不能解决缺乏可预测性而产生的各种问题。

3. 欧盟环境法中的预防原则在自动化决策当中的应用可行性分析

3.1. 数据保护影响性评估的监管手段(DPIA)的作用和局限性

在算法环境下进行自动化决策，结果的不确定性与在自然环境中进行的某些项目性质类似。这种相似源自于二者都有可能对公民个人、社会和国家产生潜在的严重后果，而后果往往难以预见。当欧盟考虑到在环境和公共健康保护领域同样存在这种不可预测性时，其往往会适用预防原则，那么该原则是否具有适用与算法环境的可能性？

在笔者看来，预防原则对环境保护义务人提出了先行处理的义务，这个义务同样适用于数据管理人。其类似于我国的比例原则，在无可预测性的情况下，必须选择损害最小的方式实施活动。GDPR 通过适用该原则，创设出了数据保护影响性评估的监管手段(DPIA)。该手段要求数据控制者对数据主体和社会

层面可能造成的高风险进行评估，内容是在启动数据处理和自动化决策之前和规定的其他特殊情况下，为数据控制者设立了对信息主体可能造成的权益损害从不同维度进行影响性评估的义务，从而帮助数据控制者尽可能对风险进行识别分析，欧盟成员国的监管机构应准备好受 DPIA 约束的各种风险处理类型清单。这表明，DPIA 作为一种工具，能够降低自动化决策带来的风险。

然而 DPIA 仍具有其局限性，根据 GDPR 规定，参与准备和讨论 DPIA 清单过程的机构仅限于非常狭窄的范围(数据控制者和各国监管机构)，数据主体仅在“适当情况下”才可参与该讨论过程，且他们并没有义务公布讨论结果。因此 DPIA 只能在很小程度上避免对数据主体可能产生的侵害。

3.2. 预防原则的概念及由来

针对这类不可预测的挑战，欧盟环境法采取的措施为适用预防原则。该原则可概括为，在缺乏关于某一种行动结果的科学证据时，有必要将保护环境的价值作为优先事项予以考虑。在国际法的发展历程中，预防原则首次于 1992 年在《马斯特里赫特条约》中得以实施，最初其仅限于关于环境采取的一系列行动，在欧盟诉比利时向另一成员国倾倒废物一案中，欧洲法院提到：“欧盟环境政策……应当以预防原则为基础[8]。”该条款为欧盟法院的若干判决提供了依据，例如 1998 年英国诉欧盟委员会一案。该原则在 2006 年的适用范围被扩大到了保护民众健康的领域，但是相关判例却早已经提出其扩张解释的需求，如美国生物制药公司 Alpharma 诉欧盟理事会一案所述：“《共同体条约》第 174 条(原第 130 条)确立了预防原则为环境法领域的基本原则，同时也适用于农业政策领域……虽然基础条约和二级立法未对预防原则的定义和内涵做出进一步规定，但是其适用符合欧盟关于提供高水平卫生保护的宗旨。动物添加剂中抗生素对人类危害存在不确定性的影响，欧盟机构可以根据预防原则采取保护性措施，无需等到现实危害变得十分明显[9]。”由于欧盟法律中预防原则的含义范围有所扩大，其实质可概括为欧洲法律中庄严载入的一般规则，在情况不确定的情况下，健康(及类似权利)的考量优先于经济利益。

3.3. 预防原则适用于公共决策的前提

从立法技术及解释技术来说，该原则或许存在适用到算法环境的可行性。该可行性建立在两个前提条件之上：第一，预防性原则直到目前的扩大解释仅针对保护公众健康的领域当中，在公共健康领域中的自动决策所产生的风险是否该当于对公共健康的物质化损害结果？例如欧洲广泛使用的 RiskER 等自动化系统，其作用是预测民众的住院风险，当该自动化决策出现风险，不但会对数据主体的数据权利产生损害，更有甚者会直接威胁到该主体的生命权和健康权。在这种情况下，应充分启动事前评估，适用预防原则，甚至直接将该类事件的人为审查优先级提至最高。风险感知活动既是一种需要运用科学理性的活动，也是一种需要运用社会理性的活动。这两种理性保持着互相交织、互相依赖的状况，从而使人们全面地把握事物的本质[10]。

第二，使预防性原则成为自动化决策早期阶段的指示性原则。自动化决策系统的前端治理具有重要意义，它能够将有些问题在自动化决策系统的开发和设计阶段就进行处理。在实践中，自动化决策系统的评估就是一种前端保护机制[11]。将设计和实施数据保护法律规定的时间点作为预防性原则应用的起点，将其作为数据保护法规和数据保护行动的指示。完善 DPIA 的适用条件，目前 DPIA 的适用条件尚显苛刻，应当细化 DPIA 各种措施的规定和具体适用场景，增加评估的形式和评估过程的透明度和包容度。在满足以上两个条件的情况下，在自动化决策中适用预防原则的可行性便会极大地增加。

4. 自动化决策在我国可能引发的风险挑战

欧盟在自动化决策上的规制思路也影响着我国关于算法自动化决策的立法建议的思考。我国许多学

者视《通用数据保护条例》为立法典范，但经过上文分析，该条例由于立法语言较为笼统，加之算法技术发展差距，我国面临的自动化决策风险有其独特之处。

第一，自动化决策会导致“信息茧房”风险。算法带来的自动化决策会导致人类决策的自主性受到限制，由于算法向人类推荐的决策选项已经为其黑箱中经过筛选的信息，随着用户使用自动决策的次数增多，在短时间内在注意力有限的情况下，用户仅能被锁定在由自动化算法推送的某类固定信息中，当用户对此类固定信息形成稳定的阅读习惯后，便再难以关注其他领域的信息或即时热点，久而久之，造成用户思维固化、盲目自信，认知结构单一，甚至模糊或淡化对现实社会的真实感知[12]。

第二，自动化决策会导致算法歧视风险。即使是算法程序的运行存在于黑箱之中，算法的设计者在初始算法之初亦以自身的使用习惯、个人偏见等带入算法，在开端便形成对特定群体可能的歧视。随着算法的自身不断完善，对用户画像、消费习惯等用户信息收集整理，为每个用户量身打造并呈现了各种信息，这既造成了用户产生“信息茧房”，也对“价格敏感型用户”或“时间敏感型用户”产生了算法歧视。除消费场景外，在金融、保险等行业进行自动化决策依然存在这些风险。

第三，自动化决策会导致用户信息泄露风险。当今世界已经进入智能时代，智能手机等电子设备不可或缺，一旦数字设备连接网络，互联网服务供应商、移动运营商、电子邮件账号供应商等自然负有保护用户在线对话、发布图片等使用痕迹的隐私信息的信息责任。但是，一旦国外警方或其他政府机构要求调取信息时，根据电子前沿基金会的报告，大多数服务供应商(如雅虎、微软、脸书和谷歌)均在未通知用户的前提下全盘提交[13]。我国尚处于个人信息保护立法的起步时期，虽然《个人信息保护法》已经公布生效，但是仍有大量信息处理者通过自动化决策导致用户信息泄露的案例发生，由于智能设备的发展过于快速，配套软件也层出不穷，自动化决策应对何种范围内的信息进行收集加工，加工的程度等也有待于法律或司法解释的细化。

5. “预防原则”对我国自动决策规制的启示

由于数字环境风险与自然环境风险治理的相似性，预防原则为数据影响性评估手段的完善和细化提供了可行性。该原则也成为了支持数据保护立法进一步发展的工具。它可以通过参与监管数字环境成为解决风险不确定性的指导方针。在欧盟，只要是涉及到公共健康政策的自动化决策，就符合当前预防原则的适用条件，但是预防原则不会改变欧盟现有的监管框架，也很难将其类推适用至其他领域，受制于其立法原则与政府体制，欧盟只能对预防原则进行有限度地扩大解释，导致无法从根源上应对各类风险挑战。

我国《个人信息保护法》第 55 条、56 条中对信息处理者的个人信息影响评估作出了类似规定。虽然在以上规定中评估的标准和方式未详细论述，但在遵循比例原则与合目的性原则的前提下，《个人信息保护法》对即将面临的自动化决策风险挑战留有余地，随时可以通过司法解释进行完善和细化，这与预防原则的内涵具有异曲同工之妙。通过分析预防原则如何在 GDPR 中发挥其作用及适用上的瓶颈，对我国完善信息方面立法也具有积极意义。

参考文献

- [1] 唐林垚.“脱离算法自动化决策权”的虚幻承诺[J]. 东方法学, 2020(6): 18-33.
- [2] 中国政法大学互联网金融法律研究院. GDPR 中文版[EB/OL]. <https://www.cnblogs.com/ostin/p/9295762.html>, 2023-03-08.
- [3] Goodman, B. and Flaxman, S. (2016) European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation”. *AI Magazine*, **38**, 50-57. <https://doi.org/10.1609/aimag.v38i3.2741>
- [4] Malgieri, G. and Comandé, G. (2017) Sensitive-by-Distance: Quasi-Health Data in the Algorithmic Era. *Information & Communications Technology Law*, **26**, 229-249. <https://doi.org/10.1080/13600834.2017.1335468>

-
- [5] Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. and Floridi, L. (2018) Artificial Intelligence and the “Good Society”: The US, EU, and UK Approach. *Science and Engineering Ethics*, **24**, 505-528. <https://doi.org/10.1007/s11948-017-9901-7>
- [6] Malgieri, G. and Comandé, G. (2017) Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. *International Data Privacy Law*, **7**, 243-265. <https://doi.org/10.1093/idpl/ix019>
- [7] Selbst, A.D. and Powles, J. (2017) Meaningful Information and the Right to Explanation. *International Data Privacy Law*, **7**, 233-242. <https://doi.org/10.1093/idpl/ix022>
- [8] European Court (1992) Commission of the European Communities v Kingdom of Belgium. <https://eur-lex.europa.eu/legal-content/EN/TEXT/?uri=CELEX:61990CJ0002>
- [9] 陈亚芸. EU 和 WTO 预防原则解释和适用比较研究[J]. 现代法学, 2012, 34(6): 146-157.
- [10] (德)乌尔里希·贝克. 风险社会[M]. 南京 译林出版社, 2004: 102-105.
- [11] 郑智航. 平衡论视角下个人免受自动化决策的法律保护[J]. 政法论丛, 2022(4): 94-105.
- [12] 孙建丽. 算法自动化决策风险的法律规制研究[J]. 法治研究, 2019(4): 108-117.
- [13] 陈颺, 裴亚楠. 算法决策风险防范的法制路径研究[J]. 重庆邮电大学学报(社会科学版), 2021, 33(3): 72-81.