

## 智能行政中自动化决策拒绝权的证成与适用

翁明杰<sup>\*</sup>

**内容提要：**随着数字技术的迭代更新，自动化决策的应用范围不断扩大、应用程度不断加深，给现代行政法治带来新的挑战。《个人信息保护法》第24条第3款规定的“自动化决策拒绝权”虽然构造相对粗糙，条文设计弹性不足，但是该规定似乎给智能行政情境下研究如何保障行政相对人合法权益提供了新方向。在智能行政中嵌入自动化决策拒绝权不仅是权力与权利不对等的必要化解路径、智能行政要求嵌入更高强度人工干预的必然举措，还与自动化决策“人在回路”的治理路径相契合。禁令路径或权利路径因不同程度上的局限被排除在自动化决策拒绝权的进路选择范围之外。折中进路以其宏观性、弱对抗性和动态性的优点，弥补其他两种路径的局限，成为自动化决策拒绝权的进路选择。在折中进路指引下，自动化决策拒绝权中的适用场域应当拓宽、“对个人利益有重大影响”的适用前提需要划定、自动化决策拒绝权的适用衔接应当优化。

**关键词：**智能行政 自动化决策拒绝权 “人在回路” 折中进路 适用规则

### 一、引言

随着大数据、人工智能、算法等数字技术的迭代更新，行政机关开始将数字技术深度应用于行政处罚、行政决策、行政许可、行政登记等行政治理领域，给传统行政活动方式带来了深刻变革。行政机关对数字技术的依赖程度不断加深，行政机关已然迈入智能行政时代。在智能行政时代，行政机关应用数字技术在提升行政效率的同时也给依法行政原则、正当程序原则、行政公开原则和行政公平原则等现代行政法治带来严重挑战。<sup>〔1〕</sup>为此，学界纷纷为应对算法、人工智能等数字技术给行政治理带来的挑战提出相应举措。

<sup>\*</sup> 翁明杰，中国政法大学法学院博士研究生。

本文为国家社会科学基金青年项目“行政审批数字化转型的法治因应研究”（24CFX017）的阶段性成果。

〔1〕 参见张恩典：《人工智能算法决策对行政法治的挑战及制度因应》，载《行政法学研究》2020年第4期。

自 20 世纪 50 年代起，算法自动化决策即开始应用于政府的公共行政。<sup>〔2〕</sup>《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）第 73 条第 2 项对自动化决策作出界定：“自动化决策，是指通过计算机程序自动分析、评估个人的行为习惯、兴趣爱好或者经济、健康、信用状况等，并进行决策的活动。”为避免智能行政中算法决策权的异化，学界从正当程序展开释义，要求行政机关建立起与行政相对人的“对话”机制，将“说明理由机制”“算法解释权”等机制嵌入算法治理中，恪守正当程序原则的底线。<sup>〔3〕</sup>《个人信息保护法》作为一部公私法交融、超越单一部门法的个人信息保护法，其中的规定自然也成为行政机关处理个人信息时应遵循的“信条”。《个人信息保护法》第 24 条第 3 款规定：“通过自动化决策方式作出对个人权益有重大影响的决定，个人有权要求个人信息处理者予以说明，并有权拒绝个人信息处理者仅通过自动化决策的方式作出决定。”该条款赋予个人对行政机关通过自动决策处理个人信息说“不”的权利，在立法上确立起“自动化决策拒绝权”。<sup>〔4〕</sup>尽管该条文规定个人可以拒绝信息处理者自动化处理个人信息，但是个人何时提出、如何提出、提出后信息处理者是必然引入人工审核还是继续“自动化”路径、个人如何救济这些问题似乎都未在该条文中得到解释。

随着智能行政的作用场域不断扩大，自动化决策拒绝权的适用场域能否不再局限于自动化行政决策的信息处理中，而进一步扩大到自动化行政决策的各个领域？《个人信息保护法》的条文能否为行政相对人对智能行政享有“自动化决策拒绝权”的论证提供新方向？这些问题都需要进一步探讨。为此，本文将从《个人信息保护法》第 24 条第 3 款出发，分析“自动化决策拒绝权”的立法构造，确认自动化决策拒绝权到底是行政相对人的“权利”还是行政机关处理个人信息的“禁令”，研究“自动化决策拒绝权”的适用范围能否扩大到行政决策的各个领域，进而分析“自动化决策拒绝权”的适用规则。

## 二、自动化决策拒绝权的规范梳理与适用困境

从现有法律规范来看，我国仅有《个人信息保护法》对自动化决策拒绝权作出规定。应当说，《个人信息保护法》的自动化决策拒绝权条款很大程度上借鉴了欧盟《通用数据保护条例》（General Data Protection Regulation，以下简称 GDPR）第 22 条第 1 款的规定。<sup>〔5〕</sup>因此，对

〔2〕 参见张凌寒：《算法自动化决策与行政正当程序制度的冲突与调和》，载《东方法学》2020 年第 6 期。

〔3〕 参见张凌寒：《算法自动化决策与行政正当程序制度的冲突与调和》，载《东方法学》2020 年第 6 期；蔡星月：《算法决策权的异化及其矫正》，载《政法论坛》2021 年第 5 期；王青斌、翁明杰：《行政处罚中智能行政裁量的实践隐忧与进路调适》，载《北京行政学院学报》2023 年第 3 期。

〔4〕 根据欧盟《通用数据保护条例》翻译不同，目前学界关于该叫法也存在区别。有的学者将其称为“完全自动化决策拒绝权”（参见王苑：《完全自动化决策拒绝权之正当性及其实现路径——以〈个人信息保护法〉第 24 条第 3 款为中心》，载《法学家》2022 年第 5 期），也有学者将其称为“免受自动化决策约束权”“免受自动化决策权”（参见蔡星月：《算法决策权的异化及其矫正》，载《政法论坛》2021 年第 5 期；张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021 年第 5 期），还有学者将其称为“脱离算法自动化决策权”（参见唐林垚：《“脱离算法自动化决策权”的虚幻承诺》，载《东方法学》2020 年第 6 期）。笔者以为，虽然各种叫法存在差异，但本质上都明确“有权拒绝个人信息处理者仅通过自动化决策的方式作出决定”。为了方便行文，本文统一将其称为“自动化决策拒绝权”。

〔5〕 GDPR 第 22 条第 1 款规定：“数据主体有权不受仅基于自动处理得出的决定的制约，以避免对个人产生法律影响或与之相类似的显著影响，该自动处理包括画像。”有学者认为 GDPR 第 22 条第 1 款规定的是“不受完全自动化决策约束权”（right to not be subject to automated decisions）。参见王苑：《完全自动化决策拒绝权之正当性及其实现路径——以〈个人信息保护法〉第 24 条第 3 款为中心》，载《法学家》2022 年第 5 期。

《个人信息保护法》中的自动化决策拒绝权的规范展开，除了需要依据现有法律规定、对比《中华人民共和国个人信息保护法（草案）》[以下简称《个人信息保护法（草案）》]，还可以参考 GDPR 的相关规定。

（一）《个人信息保护法》中自动化决策拒绝权的规范展开

1. 个人信息保护中的自动化决策拒绝权

《中华人民共和国个人信息保护法释义》指出，《个人信息保护法》第 24 条第 3 款有着两层含义：一是赋予个人对重大自动化决策的要求解释权，这是个人行使选择权和决定权的基础；二是赋予个人要求对自动化决策进行人工干预的权利。<sup>〔6〕</sup>从自动化决策拒绝权的角度来看，该条款从适用前提、适用范围、适用阶段对个人行使自动化决策拒绝权作出规定。（参见表 1）

表 1 《个人信息保护法》第 24 条第 3 款中的自动化决策拒绝权

| 适用前提       | 适用范围        | 适用阶段 |
|------------|-------------|------|
| 对个人权益有重大影响 | 仅通过自动化决策的方式 | 作出决定 |

首先，个人只有在信息处理者通过自动决策方式对个人权益有重大影响前提下才可以行使该权利。换言之，个人对信息处理者能否行使自动化决策权的判断标准是“重大影响”。至于如何界定“重大影响”，应当具体到个人信息处理的具体场景和个人的具体特征。<sup>〔7〕</sup>从理论上来看，“重大影响”的判断需要达到“可能侵害个人信息权益”的程度，“重大影响”的适用需要与一般意义上的拒绝权作区分。<sup>〔8〕</sup>

其次，自动化决策拒绝权的适用范围是“仅通过自动化决策的方式”。数字技术并非以绝对稳定的状态嵌入个人信息活动中，它会根据数字技术的发展，不断转换适用深度与场景。<sup>〔9〕</sup>有学者根据人与自动化系统的关系，将自动化分为完全自动化系统（fully automated systems）和混合自动化系统（mixed systems）。<sup>〔10〕</sup>结合《个人信息保护法》第 73 条第 2 项对“自动化决策”的规定，当信息处理者通过“混合自动化决策”方式作出决定时，个人不得行使自动化决策拒绝权。

最后，自动化决策拒绝权的适用阶段是“作出决定”而非“作出的决定”。在个人信息处理活动中，信息处理者对个人信息的处理包括处理过程和处理结果。作出决定属于单方行为，只要对个人采取具有一定程度约束力的态度/立场，或者可能根据这一态度/立场采取针对个人的活动，就可以将这一行为视为作出决定。《个人信息保护法》规定，个人认为信息处理者自动化决

〔6〕 参见杨合庆主编：《中华人民共和国个人信息保护法释义》，法律出版社 2022 年版，第 76 页。  
〔7〕 参见江必新、郭锋主编：《〈中华人民共和国个人信息保护法〉条文理解与适用》，人民法院出版社 2021 年版，第 235 页。  
〔8〕 参见江必新、李占国主编：《中华人民共和国个人信息保护法条文解读与法律适用》，中国法制出版社 2021 年版，第 88 页。  
〔9〕 参见李晴：《论自动化行政的正当性基础》，载《环球法律评论》2023 年第 1 期。  
〔10〕 See Danielle Keats Citron, *Technological Due Process*, 85 Washington University Law Review 1249, 1263 - 1264 (2008).

策可能对个人权益有重大影响，就可以说“不”，不需要考虑/判断该决定是否已经作出。此时，“决定是否已经作出无关紧要”〔11〕。

## 2. 《个人信息保护法》中自动化决策拒绝权的规范不足

应当说，自动化决策拒绝权是我国首次从法律层面对自动化决策作出的回应，进一步加强了对个人信息的法律保障。但也有学者主张，《个人信息保护法》第24条第3款的构造过于粗糙，条文设计弹性不足，需要进一步补强。〔12〕

首先，《个人信息保护法》未对适用前提中“重大影响”的判断标准作出规定。相较于《个人信息保护法（草案）》一审稿中的“个人认为自动化决策对其权益造成重大影响的”适用前提，《个人信息保护法》做了“去主观化”的处理，即删除“个人认为”的表述，保证法律规范对客观法秩序的维护。但是《个人信息保护法》对自动化决策拒绝权的判断标准仍然是一种较为模糊的立法表达。“上述判断标准有待司法机关与实践经验的进一步总结。”〔13〕“重大影响”的界定过宽，可能会妨碍自动化决策技术的良性发展；“重大影响”的界定过窄，可能会侵害个人的相关合法权益。为此，亟须立法机关/司法机关对“重大影响”的判断标准作进一步释明。

其次，自动化决策拒绝权是否存在例外情形，仍需进一步研究。〔14〕欧盟GDPR规定了自动化决策拒绝权行使的例外情形，譬如数据主体明确同意、自动化决策是数据主体与数据控制者订立等等。个人信息保护的实践纷繁复杂，是否存在例外情形仍需进一步探明，必要时可借鉴GDPR的有关规定。

## （二）智能行政中自动化决策拒绝权的适用困境

随着算法、人工智能、大数据等数字技术不断革新，行政审批、治安监控、交通执法、信用评估、风险预防等领域均受到数字技术的不断渗入。数字技术借助其深度学习、大数据计算等优势对现有的社会治理模式进行变革。〔15〕《法治政府建设实施纲要（2021—2025）》（以下简称《纲要》）指出，要“坚持运用互联网、大数据、人工智能等技术手段促进依法行政，着力实现政府治理信息化与法治化深度融合，优化革新政府治理流程和方式，大力提升法治政府建设数字化水平”，从而“全面建设数字法治政府”。全面建设数字法治政府，不仅要求国家事务和社会事务需要进行大数据归集和全流程在线通办，也要求算法技术嵌入行政的深度与广度不断加深。〔16〕行政机关越来越依赖将算法、人工智能等数字技术应用于行政活动中，以此作出行政决策。自动化行政决策的纵深发展，在给治理、服务、决策模式带来根本性改变的同时，〔17〕也因“算法黑

〔11〕 王苑：《完全自动化决策拒绝权之正当性及其实现路径——以〈个人信息保护法〉第24条第3款为中心》，载《法学家》2022年第5期，第82页。

〔12〕 参见林涇民：《自动决策算法的风险识别与区分规制》，载《比较法研究》2022年第2期。

〔13〕 江必新、李占国主编：《中华人民共和国个人信息保护法条文解读与法律适用》，中国法制出版社2021年版，第88页。

〔14〕 参见江必新、郭锋主编：《〈中华人民共和国个人信息保护法〉条文理解与适用》，人民法院出版社2021年版，第236页。

〔15〕 参见翟月荧：《行政决策中算法决策的风险识别与规制路径》，载《行政管理改革》2023年第7期。

〔16〕 参见马长山：《数字法治政府的机制再造》，载《政治与法律》2022年第11期。

〔17〕 参见张建锋：《数字政府2.0：数据智能助力治理现代化》，中信出版集团2019年版，第42页。



箱”“自动化裁量”等固有问题给行政相对人合法权益造成侵害。《个人信息保护法》第 24 条第 3 款“自动化决策拒绝权”的规定似乎给智能行政情境下研究如何保障行政相对人合法权益提供了新方向。但是，自动化决策拒绝权能否应用于整个自动化行政决策场景中、自动化决策拒绝权的性质如何界定等问题亟待进一步厘清。

#### 1. 自动化决策拒绝权进路选择不清

GDPR 第 22 条规定的自动化决策拒绝权一直存在“禁令路径”与“权利路径”的争论，给立法实践带来复杂的影响。<sup>〔18〕</sup> 尽管有学者认为《个人信息保护法》第 24 条第 3 款不同于 GDPR 模棱两可的规定，直接表达为一项权利，<sup>〔19〕</sup> 但是这种进路忽视了“以赋权与救济为核心的传统法律制度体系在面对智能社会时出现的功能危机”<sup>〔20〕</sup>。随着数字技术的迭代更新，行政相对人可能无法获知行政机关所作出的行政决策是否为“自动化决策”。毕竟，在这种情况下，算法技术完全可以不经由个人知悉的情况下直接作出行政决策。<sup>〔21〕</sup> 如果对自动化决策拒绝权性质界定采用权利进路，行政相对人还可能面临有权却不知“何时、是否应该”的权利行使困境。

如果采用禁令路径，行政机关就不能借助算法技术实施完全自动化决策。2009 年德国《联邦数据法》第 6a 节就是禁令路径的直接呈现。禁令路径的选择保护了人免受算法的滋扰与操控，表达了立法者对创新技术发展的保守态度，但是可能阻碍算法技术的良性发展与合理应用，<sup>〔22〕</sup> 还可能导致举证责任的转变——个人需要就自动化决策对其个人权益造成重大影响承担举证责任。<sup>〔23〕</sup> 从现有研究来看，在将自动化决策拒绝权应用于智能行政前，我们需要先对自动化决策拒绝权研究进路作出抉择。

#### 2. 自动化决策拒绝权适用场域不明

首先，自动化决策拒绝权全场域适用缺乏法律规范依据，有待进一步证成。行政决策是行政机关的一项行政活动，是行政机关及其公务员为了实现行政目标，依据既定政策和法律，就要解决的问题拟定并选择活动方案的行为过程。<sup>〔24〕</sup> 行政决策包括行政机关依据法律、政策作出的各类行政活动。《个人信息保护法》规定了自动化决策拒绝权，但是该条款仅适用于个人信息保护领域。行政相对人在行政机关的智能行政中能否行使该权利仍未得到理论上的证成。换言之，行政机关在行政审批、治安监控、交通执法、信用评估、风险预防等领域借助数字技术对行政相对人作出行政处罚、行政许可、行政奖励等行政行为时，行政相对人能否以行政机关自动化决策为由，拒绝行政机关基于算法等数字技术作出行政决策需要进一步论证。

其次，自动化决策拒绝权适用的决策类型能否进一步扩充需要充分澄清。当前学界对智能行政的类型划分除了前文提及的“二分法”外，还有“三分法”“四分法”。“三分法”根据人工智

〔18〕 参见张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021 年第 5 期。

〔19〕 参见吕炳斌：《论个人信息处理者的算法说明义务》，载《现代法学》2021 年第 4 期。

〔20〕 郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论丛》2022 年第 4 期，第 99 页。

〔21〕 参见郑智航：《人工智能算法的伦理危机与法律规制》，载《法律科学（西北政法大学学报）》2021 年第 1 期。

〔22〕 参见张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021 年第 5 期。

〔23〕 参见程啸：《个人信息保护法理解与适用》，中国法制出版社 2021 年版，第 223 页。

〔24〕 参见刘莘主编：《法治政府与行政决策、行政立法》，北京大学出版社 2006 年版，第 79 页。

能自我学习能力的强弱，将智能行政分为“信息收集的自动化”“流程或手续的自动化”“智能流程的自动化”。〔25〕“四分法”以数字技术的发展水平和信息处理是否需要人工干预为标准，将智能行政分为“自动化辅助行政”“部分自动化行政”“无裁量能力的完全自动化行政”“有裁量能力的完全自动化行政”。〔26〕《个人信息保护法》将有人工干预的自动化决策排除在个人行使自动化决策拒绝权的决策类型外。在自动化决策中，尽管有人的实质干预，但如果决策者存在明显偏见，仍然有可能在智能系统中嵌入有意或无意的甚至是不易被察觉的歧视。〔27〕自动化决策拒绝权适用的决策类型是否要“刻舟求剑”——局限于完全自动化行政/无裁量能力的完全自动化行政和有裁量能力的完全自动化行政中，仍需进一步研究。

### 3. 自动化决策拒绝权适用前提不明

首先，《个人信息保护法》第24条第3款对自动化决策拒绝权适用前提设定为“对个人权益有重大影响”，体现了极强的个人信息保护指向。但如果将自动化决策拒绝权适用于智能行政的全场域，自动化决策拒绝权是否只需要考虑“个人权益”还需要慎重对待。法律的终极目标不仅需要保护个人的自由和权利，还需要兼顾整个社会的发展乃至人类的生存。〔28〕在智能行政中，行政机关借助算法等数字技术开展行政活动，不仅涉及个人信息的收集、处理，还涉及对行政相对人人身自由的限制、财产的剥夺等等。因此，在确认自动化决策拒绝权适用场域的基础上，还需要对自动化决策拒绝权适用前提中的“个人权益”作进一步说明。

其次，《个人信息保护法》第24条第3款中又一关键词为“重大影响”。重大影响的范围到底有多广，本次立法未予以说明。“重大影响”的内涵与外延具有不确定性和可解释性，在适用过程中需要个案衡量。〔29〕GDPR规定了自动化决策拒绝权行使前提是“产生法律或类似的重大影响”。该前提的限制范围过大，法律影响包括对民事权利产生的影响、有行政法意义的对行政相对人产生的影响、对刑法和宪法所保护的基本权利造成的侵害，这就使得该条款在适用过程中存在巨大的困难。〔30〕在智能行政中，行政机关需要进一步结合实践考察，对“重大影响”作出科学化判断，实现“在顾及数据主体权利前提下，有效平衡个人、社会和国家之间的利益”〔31〕。

### 4. 自动化决策拒绝权适用衔接不明

在智能行政中，行政相对人的“对抗”举措缺乏体系性构建。首先，自动化决策拒绝权与其他权利之间缺乏衔接机制。智能行政在提高行政效率、便捷行政相对人的同时，也带来诸多技术、伦理、法律问题。譬如行政相对人程序权益被忽视、程序价值被压缩的问题愈发凸显，

〔25〕 参见胡敏洁：《自动化行政的法律控制》，载《行政法学研究》2019年第2期。

〔26〕 参见马颜昕：《自动化行政的分级与法律控制变革》，载《行政法学研究》2019年第1期。

〔27〕 See Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a "Right to an Explanation" is Probably Not the Remedy You are Looking For*, 16 Duke Law and Technology Review 18 (2017).

〔28〕 参见梁上上：《利益衡量论》（第3版），北京大学出版社2021年版，第89页。

〔29〕 参见程啸：《个人信息保护法理解与适用》，中国法制出版社2021年版，第233页。

〔30〕 参见蔡星月：《算法决策权的异化及其矫正》，载《政法论坛》2021年第5期。

〔31〕 郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论丛》2022年第4期，第100页。

换言之，行政相对人的程序权益在智能行政中的保障明显不足；<sup>〔32〕</sup>又譬如，在算法设计和研发的过程中，计算机专家/算法工程师可能更关注有用性和实效性，而忽视追求目的正当性以及算法活动可能带来的风险；<sup>〔33〕</sup>等等。在智能行政中，行政机关赋予行政相对人对抗算法的权利，根本目的是保障其合法权益，但是，自动化决策拒绝权并非应对自动化决策的万能权利，<sup>〔34〕</sup>它需要与其他权利共同作用方可发挥最大的效用。行政机关需要保障行政相对人在智能行政中行使“知情权”“要求说明理由权”“参与权”“监督权”“删除权”等等。GDPR 以序言第 71 条为核心展开，不以单一、个别权利为依托，而是齐头并进地构建了贯穿自动化决策事前、事中和事后的全流程联动体系，包括获得人类干预的权利、表达异议和质疑的权利、获取相关解释的权利，落实了数据保护的基本要求，充分保障了数据主体权益。<sup>〔35〕</sup>我国智能行政中行政相对人权益保障联动体系亟待进一步构建，以实现个人自动化决策拒绝权与其他权利的有效互动。

其次，自动化决策拒绝权适用效果不彰。法律规范实施效果评价，一方面需要关注法律规范本身设计的“精巧度”、与实践需求的“适配度”，另一方面需要关注法律规范的实施、适用效果。设置自动化决策拒绝权，本质上是“人工干预”的直接要求，也是机器人运行中“机器人要让位于人类”的必然规律。<sup>〔36〕</sup>人类参与到行政决策中不可流于形式，必须是实质参与，否则人类的行政参与将沦为“橡皮图章式”角色定位。<sup>〔37〕</sup>为了避免形式化参与，学者建议引入算法影响评估制度、算法审计制度，评价算法实施效果。<sup>〔38〕</sup>个人向行政机关提出拒绝应用自动化技术作出行政决策的要求后，行政机关作出的回应由谁来监督、如何开展监督、如何追责等等，这些问题当前缺乏进一步探讨，也无实践案例作为例证，需要进一步挖掘，以提升自动化决策拒绝权的适用效果。

### 三、智能行政中自动化决策拒绝权的证成

在智能行政中，个体在算法的操控下，呈现出被不断系统化和数据化的特点，逐步陷入丧失主体性的数字化生存困境。<sup>〔39〕</sup>“高级人工智能的出现可能意味着人类的灭亡。”<sup>〔40〕</sup>为有效应对

〔32〕 参见王青斌、翁明杰：《行政处罚中智能行政裁量的实践隐忧与进路调适》，载《北京行政学院学报》2023 年第 3 期，第 102 页。

〔33〕 参见林涸民：《自动决策算法的风险识别与区分规制》，载《比较法研究》2022 年第 2 期。

〔34〕 参见王怀勇、朱俊达：《自动化决策中个人拒绝权的制度困境与应对》，载《西南政法大学学报》2023 年第 1 期。

〔35〕 参见唐林垚：《“脱离算法自动化决策权”的虚幻承诺》，载《东方法学》2020 年第 6 期。

〔36〕 参见〔西班牙〕阿古斯蒂·塞里洛·马蒂内斯：《人工智能在行政法领域的影响——新科技新概念》，金园燕译，载张海斌主编：《人工智能、区块链与法治：国别区域科技与法律动态》，法律出版社 2020 年版，第 108 页。

〔37〕 See Ben Wagner, *Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-making Systems*, 11 Policy and Internet 104 (2019).

〔38〕 参见张欣：《人工智能时代的算法治理：机制与方案》，法律出版社 2022 年版，第 79-131 页。

〔39〕 参见〔瑞典〕大卫·萨普特：《被算法操控的生活：重新定义精准广告、大数据和 AI》，易文波译，湖南科学技术出版社 2020 年版，第 1-4 页。

〔40〕 〔美〕卢克·多梅尔：《人工智能：改变世界，重建未来》，赛迪研究院专家组译，中信出版集团 2016 年版，第 237 页。

这种“人让位于机器、算法技术”的悖论出现，有必要重新确立人在智能行政中的主体性定位。这一方面要求行政机关在行政活动中不能完全扮演“旁观者”的角色，也要求行政相对人在行政活动中要“实质参与”。欲纾解前文论及的自动化决策拒绝权适用困境，需要先在理论上证成自动化决策拒绝权在智能行政中有存在空间。

#### （一）必要性：智能行政发展新阶段的法治因应

##### 1. 权力与权利不对等的必要化解路径

决策是人类的专属活动，是人类运用自身特有的认知与逻辑思维能力，对一定信息和经验进行占有、分析，进而将人们的思维过程和意志行动过程进行结合的产物。决策从本质上来看，其强调的是人与人之间的一种共情心与同理心。<sup>〔41〕</sup>随着数字技术的创新发展，算法、人工智能等被应用于决策中，并逐步从辅助者、参与者转变为主导者。算法具有赋能、赋权与控制三重能力，这三重能力对被作用主体形成了一种兼具支配力、控制力和影响力的新型算法权力。<sup>〔42〕</sup>这种算法权力逐步取代行政机关，虽然提高了行政效率，有助于行政机关在更短时间作出行政决策，但是“瞬间完成的行政行为……克减了相对人事前要求告知、陈述申辩、举行听证的权利，压缩甚至取消了行政机关说明理由的义务，公众参与存在被架空的风险。算法决策程序的迅速性、自主性，一定程度钳制了个人表达的实现，引发了个人在公共行政中的‘离场’”<sup>〔43〕</sup>。传统权利义务的对应关系、平衡关系都被打破，自由、平等、公正和民主等原则都将遭遇挑战。<sup>〔44〕</sup>“公众参与是现代民主的重要形式。”<sup>〔45〕</sup>在数字技术嵌入行政决策后，权力与权利本有的“隔阂”愈拉愈大，算法权力不断扩张，公众的参与权利难以保障，公众在没有任何事前知情、发表意见、说明理由、介入的情况下就被动接受来自算法技术出具的行政决策，导致“一方权力被放大，另一方权利被削弱”<sup>〔46〕</sup>，权利与权力的“天平”严重失衡。

为实现“权力—权利”间的平衡，学界围绕如何规制算法提出了算法公开、个人数据赋权与反算法歧视等路径。<sup>〔47〕</sup>算法公开要求行政机关通过解释算法打破算法黑箱，实现“算法透明”；个人数据赋权要求对现有个人算法权利的类型与体系进行改良，确保个人权利不被算法权力蚕食；反算法歧视则要求算法本身的设计、运行“中立”，不可隐设身份歧视。这些规制路径都在算法技术嵌入行政中提出“对话机制”，但是未从根本上允许行政相对人直接向智能行政说“不”。换言之，目前算法规制路径的基本前提是智能行政应然存在，也必然存在，行政相对人必须接受人工智能、算法等数字技术以全自动或半自动方式作出的行政决策。事实上，“这种内嵌于正当程序中的算法规制本质上仍旧是针对算法的程序性控制”<sup>〔48〕</sup>。自动化决策拒绝权允许行政相对人向自动化决策说“不”，而且不仅允许行政相对人对“作出的决定”说“不”，还允许行政

〔41〕 参见郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论丛》2022年第4期。

〔42〕 参见高童非：《论自动化行政中的“算法卸责”》，载《电子政务》2023年第2期。

〔43〕 覃慧：《数字政府建设中的行政程序：变化与回应》，载《行政法学研究》2022年第4期，第151页。

〔44〕 参见马长山：《迈向数字社会的法律》，法律出版社2021年版，第60页。

〔45〕 姜明安：《行政程序法研究》，法律出版社2022年版，第170页。

〔46〕 〔德〕克劳斯·施瓦布：《第四次工业革命：转型的力量》，李菁译，中信出版社2016年版，第97页。

〔47〕 参见丁晓东：《论算法的法律规制》，载《中国社会科学》2020年第12期。

〔48〕 赵宏：《公共决策适用算法技术的规范分析与实体边界》，载《比较法研究》2023年第2期，第6页。



相对人向“作出决定”这个行为说“不”。

一方面将自动化决策拒绝权的适用阶段贯彻自动化决策的始终，另一方面不再局限于程序规则内部的“缝补”，而是直接让行政相对人站在行政决策外部作出拒绝，通过补强“权利束”的要素来缩小权利与权力间的隔阂。由此可见，引入“自动化决策拒绝权”是化解智能行政中“权力—权利”不对等的必要路径。

## 2. 嵌入更高强度人工干预的必然举措

随着技术升级，自动化决策系统中的技术构造也逐步从“人+自动化技术”转变为“全自动化技术”构造，忽视了人不受支配的主体地位。<sup>〔49〕</sup>应当说，确立自动化决策拒绝权是智能行政要求嵌入更高强度人工干预的必然举措。

首先，行政机关应当在智能行政中重申人的主体地位。算法让人陷入客体化的风险中。算法权力是从客体的角度来看待个人的，强调个人是可以被计算、预测和控制的客体。<sup>〔50〕</sup>在自动化决策中，算法系统通过“资格自动认证系统、排名算法和风险预测模型决定了哪些社区会受到治安巡查，哪些家庭能够获得所需资源，谁会被列入工作候选名单，谁又要接受欺诈调查”<sup>〔51〕</sup>，从而形成智能化、自动化的行政方式。这个过程中，专家意见、民众观点均让位于算法系统，算法决策中的权威也从“人类”转为“算法系统”。算法系统从最初的辅助者角色转变为主宰者，其通过不断更新，与人类语言思维更为贴近，最终“创造世界”，动摇人类的主体地位。“这一主体地位不仅关乎于人生活的安全，也是人类文明开展的根基，其主体地位一旦动摇，政治、经济、文化、社会制度随即分崩离析。”<sup>〔52〕</sup>人是自我利益的最好判断者，没有谁能比利益者本人更了解自己的利益。<sup>〔53〕</sup>自动化决策拒绝权的价值基础就在于人的主体性的保证。<sup>〔54〕</sup>而且，从机器和人类行为特点来看，前者更擅长重复性活动，后者则擅长复杂的、需要上下文分析的活动。<sup>〔55〕</sup>应当说，无论人工智能更新到哪一阶段，人所具有的语言思维都不可以被算法系统创设的语言思维所取代。行政决策的科学化与民主化是社会各阶层利益多元化的必然要求。<sup>〔56〕</sup>自动化决策系统只能是作出行政决策的载体，绝不可“反客为主”取代行政机关作出行政决策。<sup>〔57〕</sup>

其次，行政机关不仅需要在自动化决策系统中嵌入“人工干预机制”，还需要根据自动化决策系统发展的程度，嵌入更高强度的“人工干预机制”。确立自动化决策拒绝权，本质上是个人

〔49〕 参见郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论丛》2022年第4期。

〔50〕 See John Cheney-Lippold, *We Are Data: Algorithms and the Making of Our Digital Selves*, New York University Press, 2017, p. 141.

〔51〕 〔美〕弗吉尼亚·尤班克斯：《自动不平等：高科技如何锁定、管制和惩罚穷人》，李明倩译，商务印书馆2021年版，第3页。

〔52〕 陈姿含：《人工智能算法中的法律主体性危机》，载《法律科学（西北政法大學學報）》2019年第4期，第40页。

〔53〕 参见〔英〕弗里德利希·冯·哈耶克：《个人主义与经济秩序》，邓正来编译，复旦大学出版社2012年版，第6页。

〔54〕 参见齐延平、朱家豪：《完全自动化决策拒绝权的双重功能及实现路径》，载《国家检察官学院学报》2023年第6期。

〔55〕 See Rebecca Crotoof, Margot E. Kaminski & W. Nicholson Price II, *Humans in the Loop*, 76 Vanderbilt Law Review 429 (2023).

〔56〕 参见刘莘主编：《法治政府与行政决策、行政立法》，北京大学出版社2006年版，第149-150页。

〔57〕 参见王青斌、翁明杰：《行政处罚中智能行政裁量的实践隐忧与进路调适》，载《北京行政学院学报》2023年第3期。

主体地位在智能行政中的重申，也是个人权利与行政权力适用顺序优化整合的结果。“权利的优先性理念是一个根本要素，而在公平正义中，该理念作为公平正义观点的一种形式具有其核心作用。”<sup>〔58〕</sup> 欲凸显个人权利在权力（利）位阶排序中的地位，就需要在制度设计、规则构造中落实位阶排序的结果。前文论及，现有算法规制路径本质上是内嵌于算法技术提出的优化路径，从未站在技术之外俯瞰技术。自动化决策拒绝权站在算法、人工智能等数字技术外，允许行政相对人在权益受到或可能受到侵害时，拒绝让自动化决策系统直接作出行政决策。这种规制路径相较于其他路径而言，具有更高的强度，能达到更好的规制算法的结果。

## （二）可行性：自动化决策治理进路的转化

面对完全自动化决策的风险时，欧洲和美国选择了两条不同的进路。美国选择的是“计算中立”（computation neutrality）路径，强调人与计算机之间应当保持一定的距离，尽可能实现自动化决策的准确、公正与客观。<sup>〔59〕</sup> “计算中立”的治理路径强调应当将正当程序原则作为基础，通过程序上的补强，增强数字技术的“透明度”“公众参与度”等，最终建构起“技术正当程序”（technological due process）框架。<sup>〔60〕</sup> 相反，欧洲认为机器学习、自动驾驶、自动化武器及自动化决策等人工智能运用场景中要求人类介入，即所谓的“人在回路”（human-in-the-loop）。<sup>〔61〕</sup> 换言之，即使数字技术已经发展到足够成熟的阶段，完全无需依靠人为介入也可以得出结果，“人在回路”还是要求人类作为把关者。当然，人的作用不仅局限于自动化行政的结果中，也包括在自动化行政前、自动化行政的过程中，对决策的全程产生“实质影响”。欧洲选择的“人在回路”的治理路径在其立法中有所体现。譬如《德国联邦程序法》第 35a 条和《德国联邦数据法》第 6a 节对完全自动行政持禁止态度，GDPR 第 22 条“完全自动化行政决策拒绝权”也采用“人在回路”的立法路径。<sup>〔62〕</sup>

在“计算中立”的治理路径中，智能行政中的“有裁量能力的完全自动化行政”完全有可能实现。在“人在回路”的自动化决策治理路径中，“有裁量能力的完全自动化行政”就无法实现，最多只能实现“混合自动化行政”，因为“人在回路”要求人在任何一种自动化决策中都要发挥实质作用。“人在回路”治理路径的基本假定就是，人和机器都有各自擅长的事情，在混合自动化行政决策中，二者可以达到一种最佳状态。<sup>〔63〕</sup> 随着欧盟模式的全球化推广，“人在回路”的治理路径也逐渐渗透到各国的立法、制度设计中。我国的《个人信息保护法》《中华人民共和国电子商务法》《互联网信息服务算法推荐管理规定》等法律规范中都选择“人在回路”的算法治理路径。“人在回路”面临着“一味强调‘人的干预’，却忽略人在何时干预、能否干预、干预效果、如何干

〔58〕〔美〕约翰·罗尔斯：《政治自由主义》，万俊人译，译林出版社 2000 年版，第 184 页。

〔59〕参见王苑：《完全自动化决策拒绝权之正当性及其实现路径——以〈个人信息保护法〉第 24 条第 3 款为中心》，载《法学家》2022 年第 5 期；陈悦：《论自动化行政中算法决策风险的“人在回路”治理模式》，载《行政法学研究》2024 年第 4 期。

〔60〕See Danielle Keats Citron, *Technological Due Process*, 85 Washington University Law Review 1249 (2008).

〔61〕See Meg Leta Jones, *The Right to A Human in the Loop: Political Constructions of Computer Automation and Personhood*, 47 Social Studies of Science 216 (2017); 郑戈：《人工智能与法律的未來再思考》，载《数字法治》2023 年第 3 期。

〔62〕参见王苑：《完全自动化决策拒绝权之正当性及其实现路径——以〈个人信息保护法〉第 24 条第 3 款为中心》，载《法学家》2022 年第 5 期。

〔63〕参见郑戈：《人工智能与法律的未來再思考》，载《数字法治》2023 年第 3 期。

预等技术操作问题”的诘问，也有学者主张，“计算中立”与“人在回路”并非互相替代的治理路径，二者对应智能行政发展的不同阶段。从人工智能、算法、大数据等数字技术的发展来看，数字技术的高速发展已然超出预测范围。2022 年，OpenAI 发布的深度学习技术的大型自然语言处理模型 ChatGPT “通过学习大量的自然语言文本数据，利用统计学方法和概率分布根据先前的输入预测下一个可能出现的单词或句子，生成流畅连贯、语法正确、逻辑通顺的文本”<sup>〔64〕</sup>。随着这种数字技术的应用，智能行政可能将步入“有裁量能力的完全自动化行政”阶段，“算法黑箱”“算法歧视”变得更为隐蔽，或许无法再通过“计算中立”路径中的“程序上补强”来达成治理效果。这时候就需要引入“人在回路”的治理路径。尽管该路径也有缺陷，但是可以通过制度设计、优化来发挥更好的治理效果。

智能行政时代，选择“人在回路”的数字技术治理路径已是大势所趋。自动化决策拒绝权区别于要求说明理由、知情权等程序性权利，其契合“人在回路”的治理路径，强调自动化决策中的“高强度人工干预”。因此，在智能行政中确立自动化决策拒绝权具有可行性。

#### 四、智能行政中自动化决策拒绝权的进路选择

前文对自动化决策拒绝权能否应用于智能行政中从正当性与可行性两方面展开了论证。下一步，需要确定自动化决策拒绝权的研究进路。

##### （一）自动化决策拒绝权的进路选择：禁令进路抑或权利进路？

当前学界围绕自动化决策拒绝权的进路选择到底是权利进路还是禁令进路展开讨论。这背后折射出不同的利益择向，也直接影响了自动化决策制度构建的具体编排。

禁令进路强调“禁止完全自动化决策”，即禁止行政机关对行政相对人实施完全自动化决策，除非这种决策属于法律上明确的例外情形。由此可见，禁令进路折射出立法者对自动化决策持以相对抵触和谨慎的态度。<sup>〔65〕</sup>这种进路以“禁止自动化决策为主、例外适用为辅”为原则，本质上是“事前预防”理念的落实，从根源上预防自动化决策对行政相对人权益的侵犯。有学者从 GDPR “解释指南”出发，认为 GDPR 第 22 条是“一般性禁止”，而非个人的“反对权利”。<sup>〔66〕</sup>GDPR 认为有人工干预/参与的行政决策因技术和理性避免了完全机器决策对个人尊严的减损，所以个人不得对这类决策行使自动化决策拒绝权。<sup>〔67〕</sup>

权利进路强调“赋权”，即赋予行政相对人免受自动化决策侵扰的权利。该种进路变相承认了完全自动化决策的存在，即行政机关在行政相对人拒绝前，仍可以通过自动化决策的方式对行政相对人作出行政决策，也强调行政相对人可以通过事后补救达成保护自我权益的目的。权利进

〔64〕 张欣：《面向产业链的治理：人工智能生成内容的技术机理与治理逻辑》，载《行政法学研究》2023 年第 6 期，第 44 页。

〔65〕 参见郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论丛》2022 年第 4 期。

〔66〕 参见解正山：《算法决策规制——以算法“解释权”为中心》，载《现代法学》2020 年第 1 期。

〔67〕 参见张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021 年第 5 期。

路事实上是在事前赋予行政相对人针对自动化决策权的反对权，克服了事前与事中预警机制的不足，与自动化决策运行中的知情权、表达权、提出异议权和自动决策运行后的救济权共同构筑起一套算法风险防范机制。<sup>〔68〕</sup>譬如，有学者就主张，GDPR里的自动化决策拒绝权并非完全的“禁令”，它强调一种更高的人工介入程度，与“获解释权”“被遗忘权”同处于自动化决策的个人权利体系中，三者是层层递进的关系。<sup>〔69〕</sup>还有学者从GDPR的体例编排出发，认为欧盟将自动化决策拒绝权放在权利主体权利的章节中，是权利进路。<sup>〔70〕</sup>

## （二）GDPR中自动化决策拒绝权进路选择的镜鉴

事实上，GDPR第22条自动化决策拒绝权到底采取何种立法进路一直存在争议，也影响了具体实践。除了前述提及的“禁令进路”和“权利进路”，近来有学者主张GDPR第22条本质上采取“严格禁令”与“积极权利”路径的中间路线。欧盟的立法者着眼于全要素、全过程、多场景，面向数据流生命周期建立起了“知情、参与、反对、救济”的一体化行权方案，形成了多层次、立体化的程序性联动保护机制，从广度和深度两个层面为免受自动化决策约束权奠定了配套机制。遗憾的是，目前该条款设计有着限制过多、指引不明的缺陷。<sup>〔71〕</sup>

折中进路认为，GDPR第22条既可以解释为禁令，也可以解释为一种反对权，这两种解释主要区别在于数据主体是否需要采取措施来限制自动化决策。<sup>〔72〕</sup>一方面，折中进路淡化自动化决策拒绝权的权利属性，将其从权利保护的关系网中抽离出来，从数字技术的发展对行政相对人权益的影响出发，要求个人可以主动干预，提出拒绝适用自动化决策作出行政决策的要求。相反，采取权利进路，无论是程序控制还是内嵌实体赋权，所缺失的都是在源头处对算法进入公共决策予以严格把控。<sup>〔73〕</sup>另一方面，折中进路又意识到对数字技术持以过分保守态度，不利于人工智能、大数据、算法的进一步发展，允许行政机关在征求行政相对人同意且不存在违法的情况下，借助自动化决策系统作出行政决策。这种进路克服了禁令进路和权利进路的“偏执观”和“绝对性”，实现“权利保护与产业发展间的精细平衡”<sup>〔74〕</sup>。

## （三）自动化决策拒绝权的进路选择：一种折中的进路

不可否认，自动化决策依托算法技术的决策系统替代人类决策，能够大幅提高行政效率、克服执法的恣意、保证决策标准的统一。<sup>〔75〕</sup>但是，在数字化技术的深度应用中，作为规则的代码对技术产生了规范与约束，<sup>〔76〕</sup>个人权利也会为数字技术所绑架。前文提及的禁令进路反映出立

〔68〕 See Aziz Z. Huq, *A Right to a Human Decision*, 106 Virginia Law Review 611, 611–615 (2020).

〔69〕 参见唐林垚：《“脱离算法自动化决策权”的虚幻承诺》，载《东方法学》2020年第6期。

〔70〕 参见孙跃元：《〈通用数据保护条例〉框架内的自动化决策与数据保护》，载《网络安全技术与应用》2020年第9期。

〔71〕 参见张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021年第5期。

〔72〕 See Sandra Wachter, Brent Mittelstadt & Luciano Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 International Data Privacy Law 76 (2017).

〔73〕 参见赵宏：《公共决策适用算法技术的规范分析与实体边界》，载《比较法研究》2023年第2期。

〔74〕 王苑：《完全自动化决策拒绝权之正当性及其实现路径——以〈个人信息保护法〉第24条第3款为中心》，载《法学家》2022年第5期，第80页。

〔75〕 参见马长山：《数字法治政府的机制再造》，载《政治与法律》2022年第11期。

〔76〕 参见〔美〕劳伦斯·莱斯格：《代码2.0：网络空间中的法律》（修订版），李旭、沈伟伟译，清华大学出版社2018年版，第6页。



法者对数字技术的隐忧，推崇“以人为本”，担心算法会蚕食个人权益。自动化决策是一种借助公式和程序计算获得的机器判断，缺乏可识别/解释的规则作为依据，可谓自治系统。<sup>〔77〕</sup> 自动化决策算法容易与“人本主义”相背离，<sup>〔78〕</sup> 一旦不当应用，就会祸害无穷。但是，一开始就对数字技术持有抵触情绪，会直接影响数字技术的良性发展和合理应用。权利进路反映出立法者对数字技术抱有相对乐观的态度，坚信“计算中立”。可是，算法的“原罪”在于其设计之初有意或无意地引入了某种简化和误差，而这在后续的更新迭代中被不可避免地放大。<sup>〔79〕</sup> 而且，“在双方力量悬殊时，单纯的赋权并不足以改变权力/权利之间的势差”<sup>〔80〕</sup>。由此可见，无论是禁令进路还是权利进路，都有一定的局限。GDPR 中的折中进路避免智能行政对个人权益的保护陷入两种极端，即过分关注个人权益忽视信息技术发展的极端、过分关注信息技术发展忽视个人权益的极端。

我国智能化行政可以选择折中进路，从根本上避免智能化行政对相对人权益保护陷入两种极端。折中进路并非权利进路或禁令进路在不同场景下的选择，而是具有普通适用性的自动化决策拒绝权进路选择。我国要在智能行政中构建起自动化决策拒绝权的制度体系，就需要恪守折中进路中的“三性”，即宏观性、弱对抗性和动态性，达成“一方面建立即时矫正异化的‘束状权力结构’，另一方面扩充长效制衡权力的‘算法权利版图’”<sup>〔81〕</sup> 的效果，真正发挥自动化决策拒绝权的作用。

首先，自动化决策拒绝权制度体系的构建需要遵循折中进路中的宏观性。传统禁令进路与权利进路都将自动化决策权作为制度构建的核心范畴，却从未站在自动化决策拒绝权应用于智能行政的场景来审视。这一方面导致自动化决策拒绝权闭门造车，停留在权利谱系中不停演绎。毕竟，“权利总是个别且列举性的，难以将公法中的权限和程序所具有的制度性自我规制，以及对抽象的公民整体权利保护全部归纳入个人的权利体系中”<sup>〔82〕</sup>。另一方面导致自动化决策拒绝权的研究停留在对完全自动化系统和混合自动化系统二元应用场景的“圈地运动”中。学界因禁令进路的影响，对混合自动化系统持包容态度，而对完全自动化系统赶尽杀绝。自动化行政决策的最终发展阶段应当是应用完全自动化系统的阶段。无限的“圈地运动”，只会将完全自动化行政决策阻挡在算法技术发展的大门外。因此，自动化决策拒绝权的折中进路就要求我们要在“算法决策大规模进入公共领域这一现实背景下”<sup>〔83〕</sup> 构建制度，确保研究视角的宏观性。

其次，自动化决策拒绝权制度体系的构建需要遵循折中进路中的弱对抗性。无论是禁令进路

〔77〕 See Makoto Cheng Hong & Choon Kuen Hui, *Towards a Digital Government: Reflections on Automated Decision-making and the Principles of Administrative Justice*, 31 Singapore Academy of Law Journal 875, 878 (2019).

〔78〕 参见侯东德、张可法：《算法自动化决策的属性、场域与风险规制》，载《学术研究》2022 年第 8 期。

〔79〕 参见董春雨、薛永红：《大数据哲学：从机器崛起认识到认识论方法的变革》，中国社会科学出版社 2021 年版，第 53 页。

〔80〕 赵宏：《公共决策适用算法技术的规范分析与实体边界》，载《比较法研究》2023 年第 2 期，第 6 页。

〔81〕 蔡星月：《算法决策权的异化及其矫正》，载《政法论坛》2021 年第 5 期，第 30 页。

〔82〕 〔日〕小早川光郎：《行政诉讼的构造分析》，王天华译，中国政法大学出版社 2014 年版，第 189 页。

〔83〕 郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论坛》2022 年第 4 期，第 100 页。

还是权利进路，都习惯将算法技术或个人权利中的某一方逼到“死角”。事实上，将自动化决策与人工行政决策的运行逻辑对比来看，算法决策遵循数理逻辑，法律决策依据社会逻辑，二者处于完全不同的逻辑系统，算法自动化决策仅能做到提高效率、减少工作重复和简单的推理，若要将算法决策转化为法律决策，摆在前面的技术障碍和理论阻却难以突破。<sup>〔84〕</sup>既然算法决策与法律决策属于两种不同的逻辑系统，我们就更不应当将它们放置在刻度尺的两端，相反，我们应弱化二者的对抗性，挖掘各自的优势，实现个人、社会和国家利益间的平衡，实现个人利益保护与国家数据产业发展间的动态平衡。

最后，自动化决策拒绝权制度体系的构建需要遵循折中进路中的动态性。禁令进路强调行政相对人可以在自动化决策前就拒绝行政机关借助自动化系统作出决策。权利进路克服禁令进路的局限性，尝试丰富权利谱系，将自动化决策拒绝权适用于自动化决策的全过程。尽管权利进路相较于禁令进路，关注到造成“权利—权力”失衡的某一方面原因，但是该进路过份强调自动化决策拒绝权的适用，反而导致权利行使的形式化。折中进路的动态性本质上要求自动化决策拒绝权在制度构建时增强制度弹性。一方面要求自动化决策拒绝权在不同适用场景中根据需要选择不同强度的人工干预，另一方面则要求自动化决策拒绝权在适用某一场景时，“建立特定时间段内的流动性平衡，而非某个时间点上的时时平衡”<sup>〔85〕</sup>。

## 五、智能行政中自动化决策拒绝权的适用展开

自动化决策拒绝权在适用中面临着进路选择不清，适用场域、适用前提、适用衔接不明等困境。我国应当以折中进路作为自动化决策拒绝权的进路选择，选择折中进路有助于克服禁令进路和权利进路的不足，还有助于将折中进路的宏观性、弱对抗性和动态性应用于智能行政自动化决策拒绝权适用困境的化解中。在折中进路的指引下，自动化决策拒绝权的适用场域、适用前提和适用衔接能进一步明晰。

### （一）拓宽自动化决策拒绝权的适用场域

当下自动化决策拒绝权适用场域面临的问题在于既无法确定其能否完全应用于智能行政的各个场域，也无法确定其是否能够应用于完全自动化决策中。

首先，自动化决策拒绝权应当适用于智能行政的各个场域。尽管《个人信息保护法》是一部横跨公私两域的法律，但是该法只对处理个人信息的决策活动作出规定，未涉及行政决策的其他领域。在传统行政活动中，信息资源匮乏、信息收集缓慢且成本高昂、信息处理能力羸弱。<sup>〔86〕</sup>随着科技的发展，我国的“物理政府”也因为互联网、大数据、算法、人工智能等数字技术的促

〔84〕 参见侯东德、张可法：《算法自动化决策的属性、场域与风险规制》，载《学术研究》2022年第8期。

〔85〕 蔡星月：《算法决策权的异化及其矫正》，载《政法论坛》2021年第5期，第29页。

〔86〕 参见〔英〕迈尔·舍恩伯格、肯尼思·库克耶：《大数据时代：生活、工作与思维的大变革》，盛杨燕、周涛译，浙江人民出版社2013年版，第17页。

进逐渐向“数字政府”转型。<sup>〔87〕</sup>《纲要》提出的“全面建设数字法治政府”，更是将“数字政府”的建设提到更高的层面，即要实现法治与数字的深度耦合、法治政府的数字化转型。运用算法等数字技术进行自动化决策是数字政府的核心。<sup>〔88〕</sup>自动化决策不仅包括个人信息的收集，还包括运用算法进行模型构建、风险预测、网上流程管理、网上审批等等。自动化决策拒绝权相较于以往的“说明理由”“听取意见”等程序权利具有“更强人工干预”的特点。虽然个人信息的保护强度直接影响个人利益的保障程度，但智能行政的法治化程度更是直接影响个人利益、社会利益和公共利益，还可能影响现代法治的发展程度和升级限度。因此，有必要在智能行政中引入自动化决策拒绝权，调适“权力—权利”失衡的局面，营造良好数字生态。明确自动化决策拒绝权的广泛适用场域与折中进路中的宏观性相契合。宏观性要求我们对自动化决策权的研究站在智能行政的场景来审视，实现智能行政应用情景与自动化决策拒绝权的深度融合。

其次，自动化决策拒绝权不仅可以在混合自动化系统中适用，还可以在完全自动化系统中适用。折中进路的宏观性要求自动化决策拒绝权的适用摆脱智能行政的“二分法”“三分法”乃至“四分法”的局限，以宏观视角剖析智能行政，没有必要再对自动化决策拒绝权“进行基于数据库编码的计算机自动化、基于机器学习的算法自动化和基于神经网络的超级自动化的场景划分”<sup>〔89〕</sup>。一方面，完全自动化系统可能一开始就携带“偏见、黑箱、高对话成本”，人在自动化决策的设计中，可能有意无意地将自身的价值追求、预期目标置入算法中。<sup>〔90〕</sup>以算法歧视为例，如果一开始嵌入自动化系统的大量训练数据就隐藏歧视，那无论智能系统如何更新迭代，算法歧视总会贯彻始终。另一方面，混合自动化系统可能早已成为完全自动化系统的外套。考虑到“人工智能算法在语言和思维层面愈来愈逼近人类”<sup>〔91〕</sup>，行政机关的工作人员为减少自身思考，假借“尽到人工干预义务”之名，忽略“个别情况考虑义务”<sup>〔92〕</sup>，将决策权全部托付给自动化决策系统。毕竟在算法等数字技术高歌猛进的当下，行政机关借助算法技术形成形式上的人工干预来规避“个人干预义务”的行使也非难事，但是这种人工干预对自动化决策不产生实质性影响。在混合自动化系统与完全自动化系统之间界限愈发模糊的同时，个人对自动化决策也开始愈发秉持一种“不信任”的态度。如果个人一开始就对决策主体持怀疑态度、缺乏基本信任，那无论决策主体如何公开信息都会被排斥、决策结果如何合理都会被反对，决策也无法被执行，更有甚者，决策的情境可能都无法启动。<sup>〔93〕</sup>

《个人信息保护法》第 23 条第 4 款中的“仅通过自动化决策的方式”将自动化决策拒绝权适

〔87〕 参见张建锋编著：《数字政府 2.0：数据智能助力治理现代化》，中信出版集团 2019 年版，第 40—45 页。

〔88〕 参见马长山：《数字法治政府的机制再造》，载《政治与法律》2022 年第 11 期。

〔89〕 唐林垚：《“脱离算法自动化决策权”的虚幻承诺》，载《东方法学》2020 年第 6 期，第 29 页。

〔90〕 See Pat O'Malley, *The Uncertain Promise of Risk*, 37 the Australian and New Zealand Journal of Criminology 323, 323—330 (2004).

〔91〕 郑智航：《人工智能算法的伦理危机与法律规制》，载《法律科学（西北政法大学学报）》2021 年第 1 期，第 18 页。

〔92〕 王青斌、翁明杰：《行政处罚中智能行政裁量的实践隐忧与进路调适》，载《北京行政学院学报》2023 年第 3 期，第 103—104 页。

〔93〕 参见韩春晖：《行政决策的多元困局及其立法应对》，载《政法论坛》2016 年第 3 期。

用场域限定于“自动化决策的方式”，但未限定为“完全自动化决策的方式”。本文不再拘泥于自动化决策方式的程度分类，而是根据自动化决策拒绝权适用的宏观性主张自动化决策拒绝权的适用场域应当拓宽至智能行政的各个场域。该种在折中进路指引下的规范阐释并不会与《个人信息保护法》现有规定相龃龉。因此，出于对自动化系统运行现状的考量和个人利益、社会利益和国家利益保护的需求，自动化决策拒绝权应当全场景适用于智能行政中，且无需限定仅适用于完全自动化行政中，需要全面拓宽其适用场域。

## （二）划定自动化决策拒绝权的适用前提

《个人信息保护法》第24条第3款规定自动化决策拒绝权的适用前提是“对个人权益有重大影响”。但是，该规定存在适用模糊的困境。唯一可以确定的是，自动化决策拒绝权中的“个人”必须是智能行政中享有权益的行政相对人，而不是他人。<sup>〔94〕</sup>“对个人利益有重大影响”的判断还需要在折中进路的引导下进一步展开，避免对“对个人权益有重大影响”作扩大、泛化解释，导致行政相对人滥用自动化决策拒绝权。

首先，应当将对“对个人权益有重大影响”的判断置于一个动态的分析框架内。一方面，可以通过“列举+兜底”<sup>〔95〕</sup>或“分级排序”<sup>〔96〕</sup>的方式界定“重大影响”中的判断要素。GDPR第22条将适用前提规定为“产生法律或类似的重大影响”，欧盟第29条数据保护小组认为，重大影响是对数据主体产生长期或永久影响或在最极端的情况下导致对个人的排斥或歧视。<sup>〔97〕</sup>我们可以从中提炼出重大影响的几个共同要素，具体包括法律地位的失衡、影响的持续性、影响的极端性、合法权益的明显侵害等等。这些要素可以作为列举的要素。另一方面，重大影响的判断标准需要有兜底情形，避免“判断彻底量化”<sup>〔98〕</sup>情形的出现。算法、人工智能、大数据等数字技术日新月异，对个人权益产生的影响也处在不断变化中。譬如，新出现的生成式人工智能技术就塑造了人机共生的关系，区别于深度合成技术、智能交互技术、高风险人工智能。<sup>〔99〕</sup>这种技术应用于自动化决策中，“对个人权益有重大影响”的判断标准可能又有变化。如果直接以全面列举的方式对适用前提作出界定，会难以适应实践的需求，影响判断标准的实践性。折中进路的动态性不仅要求权利谱系的动态变化，也要求自动化决策拒绝权制度设计应富有制度弹性，即动态、流动地满足实践需要。因此，重大影响的判断标准还需要设计兜底情形，即“一般理性人”<sup>〔100〕</sup>融合具体场景综合考量、合理评估后认为属于重大影响的。如此一来，对“对个人权益有重大影响”的判断就既有直接指向性，避免了模糊性缺陷，又有灵活性，能较好地适应智能行

〔94〕 参见杨立新、赵鑫：《利用个人信息自动化决策的知情同意规则及保障——以个性化广告为视角解读〈个人信息保护法〉第24条规定》，载《法律适用》2021年第10期。

〔95〕 参见王怀勇、朱俊达：《自动化决策中个人拒绝权的制度困境与应对》，载《西南政法大学学报》2023年第1期。

〔96〕 参见郑智航：《平衡论视角下个人免受自动化决策的法律保护》，载《政法论丛》2022年第4期。

〔97〕 See Article 29 Data Protection Working Party, Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation 2016/679, WP251 rev. 01, 6 February 2018, p. 21.

〔98〕 张新宝：《〈中华人民共和国个人信息保护法〉释义》，人民出版社2021年版，第198页。

〔99〕 参见韩旭至：《生成式人工智能治理的逻辑更新与路径优化——以人机关系为视角》，载《行政法学研究》2023年第6期。

〔100〕 “一般理性人”标准是指以“一般理性人”的立场融合具体场景综合考量、合理评估。参见张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021年第5期。



政的发展变化。

其次,“对个人权益有重大影响”中的“影响”既包括已经发生的影响,也包括未发生的影响。重大影响的判断本质上涉及行政裁量,需要行政机关根据影响发生的阶段、影响范围等要素动态调适判断标准的松紧:对已经发生的影响采用相对“宽松”的判断标准,对未发生的影响采用相对“严格”的判断标准。如果对未发生的影响采用过于宽松的标准,有“矫枉过正”之嫌,无益于数字技术、产业的健康发展。传统行政法学研究恪守依法行政原则,着眼于行政行为的病理性分析,缺乏“最佳行政”的分析视角。<sup>[101]</sup>新的行政法的二维结构应当是“合法行政”与“最佳行政”的二维互动。<sup>[102]</sup>“合法行政”与“最佳行政”的互动强调行政的合理性,亦对行政机关与相对人之间的关系提出更高的要求,即高权行政活动的疆域应适当缩小。对“对个人权益有重大影响”的判断既要遵循法律规范的基本要求,还要考虑判断背后的“利益博弈”,强调行政机关与相对人之间的弱对抗性,寻求最佳方案,实现内在利益的动态平衡,达成最佳行政的目标。

### (三) 优化自动化决策拒绝权的适用衔接

自动化决策拒绝权在全场域适用于智能行政后,还需要考虑如何实现自动化决策拒绝权的适用协调,构建起一个封闭周密的行政相对人合法权益保护网,改善“权力—权利”失衡的局面,弱化个人与行政机关之间的对抗性,充分发挥自动化系统的优势,推进数字技术产业的良性发展。优化自动化决策拒绝权的适用衔接,需要从自动化决策拒绝权的适用阶段和自动化决策拒绝权与其他权利有序衔接两方面展开(参见图1)。

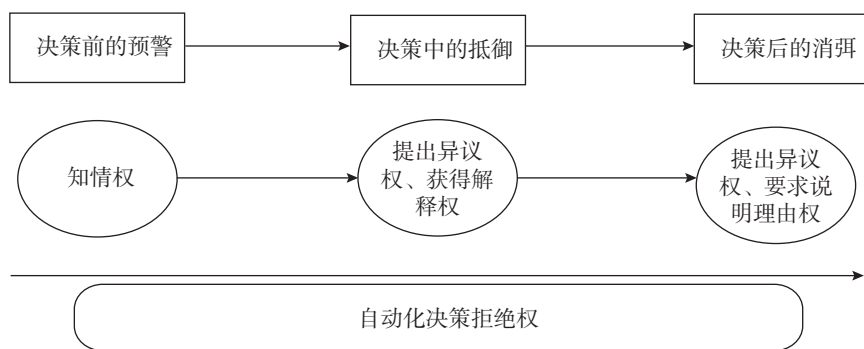


图1 自动化决策拒绝权适用衔接图

#### 1. 确立自动化决策拒绝权适用阶段

第一,现有法律规范为自动化决策拒绝权适用阶段的确立提供了依据。《个人信息保护法》第24条第3款将自动化决策拒绝权的适用阶段规定为“作出决定”。前文提及,“作出决定”区别于“作出的决定”:前者强调行政机关借助自动化系统作出决策这一活动,更多的是行为上的宣示;后者强调行政机关借助自动化系统作出的决策,更关注“最终的决策”。因此,“作

[101] 参见陈锦波:《自动化行政合法性证成的基本逻辑》,载《法学论坛》2023年第3期。

[102] 参见朱新力、唐明良等:《行政法基础理论改革的基本图谱:“合法性”与“最佳性”二维结构的展开路径》,法律出版社2013年版,第15-16页。

出决定”不在乎决定是否已经作出，在乎的是行为会对个人产生的重大影响。而且《个人信息保护法》第24条第3款的“并”字说明“要求说明理由权”与“自动化决策拒绝权”之间没有行使的前后顺序。个人有权要求信息处理者对“作出的决定”说明理由，这在实践上表示为作出决定后的说明。<sup>〔103〕</sup>而自动化决策拒绝权不要求个人必须在信息处理者作出决定后才可行使。

第二，自动化决策拒绝权可以在自动化决策的任何一个阶段提出是自动化决策“人在回路”治理路径的要求，是个人自决在数字时代的内在要求。个人自决指的是，个人在面对某种可能性的时候，有作出从事或放弃决定的自由。<sup>〔104〕</sup>在数字时代，“人可能将技术理性发展得非常完美，但其总会有所保留，而这个保留的东西就是人文”<sup>〔105〕</sup>。这种“人文关怀”应当贯彻自动化决策的始终，不应针对个人的拒绝设定不合理的“门槛”；这种“人文关怀”也不应当有行使阶段的限制，行政机关应保护个人在认为有“某种可能性”存在时就行使自动化决策拒绝权的自由。

第三，自动化决策拒绝权的全阶段适用是折中进路中宏观性的基本要求。折中进路要求我们对自动化决策拒绝权的解释不再停留在权利谱系中不停演绎。我们需要站在公民知情权、提出异议权、获得解释权等权利的外围，瞰览智能行政中的权利谱系。自动化决策拒绝权应当与知情权、提出异议权等权利相区别，不拘泥于适用阶段片段性。

## 2. 构建自动化决策中的全流程行权机制

行政机关作出自动化决策可以分为作出前、作出中和作出后。为保障行政相对人在自动化决策中的合法权益，有必要在自动化决策的每一阶段嵌入行政相对人的行权路径、确定每一阶段行政相对人的权利内容，构建起决策前的风险识别预警机制、决策中的风险抵御机制和决策后的风险消弭机制。<sup>〔106〕</sup>

在决策前的风险识别预警机制中，行政机关需要保证算法决策的公开、保证个人知情权的实现。在自动化决策系统中，行政机关会预先向其置入程序语言，组建起模拟人类思维模式的模型。<sup>〔107〕</sup>囿于算法语言的理解门槛和自我的考量，行政机关一般不愿意向公众公开算法决策内部的利益权衡、规则内容，从而会加大公众与行政机关的“代沟”。这本质上违反了行政公开原则。行政机关对“算法作出决策的基本规则和因素权重，应该对公众与社会公开”<sup>〔108〕</sup>。从行政相对人的角度而言，其可以直接以行使“知情权”为由，要求行政机关主动以一般人可以理解的表达方式公开自动化决策系统内部的规则、因素权重。

在决策中的风险抵御机制中，行政机关应当允许公众对决策系统、决策提出异议，也应当保

〔103〕 参见吕炳斌：《论个人信息处理者的算法说明义务》，载《现代法学》2021年第4期。

〔104〕 参见谢远扬：《个人信息的私法保护》，中国法制出版社2016年版，第103页。

〔105〕 〔德〕阿图尔·考夫曼：《法律哲学》（第2版），刘幸义等译，法律出版社2011年版，第143页。为避免理解上的偏差，本文将“人文”概念的表述统一为“人文关怀”，强调自动化决策拒绝权是对“人”的主体性的关照。

〔106〕 参见张欣：《免受自动化决策约束权的制度逻辑与本土构建》，载《华东政法大学学报》2021年第5期。

〔107〕 参见王青斌、翁明杰：《行政处罚中智能行政裁量的实践隐忧与进路调适》，载《北京行政学院学报》2023年第3期。

〔108〕 张凌寒：《算法自动化决策与行政正当程序制度的冲突与调和》，载《东方法学》2020年第6期，第12页。

证行政相对人在自动化决策中实现“获得解释权”。个人在发现自动化决策系统中的算法存在歧视、利益衡量存在不公的时候就可以直接向行政机关提出异议。在传统行政决策中，行政机关通过专家参与和公众参与提高决策的科学性与民主性。在自动化决策中，专家和公众在“黯然离场”中面临着权益受到侵害的风险。因此，个人可以对之前公开的决策基本规则和因素权重中存在的问题提出异议，调整其在自动化决策中的弱势地位。这也算智能时代专家意见的主导地位让位于数据与算法的补救措施。<sup>〔109〕</sup>“获得解释权”可以理解为个人权利中的“权利束”，更多应用于个人信息保护领域。<sup>〔110〕</sup>“获得解释权”本质上也是行政相对人提出异议权的延伸，可以纾解自动化决策中算法妨害等问题。<sup>〔111〕</sup>

在决策后的风险消弭机制中，行政机关既可以允许行政相对人对行政决策提出异议、寻求权利上的救济，也应当对自动化决策系统中裁量事实、裁量结果、裁量事实与裁量结果之间的因果关系进行说明。《个人信息保护法》第 24 条第 3 款对“要求说明理由权”作出了规定，明确行政相对人在自动化决策对个人权益造成重大影响时可以要求行政机关予以说明。此时，这种说明理由就成为行政机关的一项义务，应当在“对个人权益有重大影响”的适用前提下严格遵循。

在构建决策前风险预警机制、决策中风险抵御机制和决策后风险消弭机制后，还需要将自动化决策拒绝权嵌入自动化决策的各个阶段，实现自动化决策中权利的有效联动，搭建起一个立体化、全面化、动态化的行政相对人行权体系。

## 六、结 语

在当前的智能行政中，虽然有裁量能力的完全自动化行政暂无实例，但是随着大数据、人工智能、算法等数字技术的高速发展、迭代更新，有裁量能力的完全自动化行政也将出现在不久的将来。“人工智能算法独特的运行逻辑导致法律赖以生成与存在的社会结构性场景发生了重大变化，加剧了决策者与相对人之间的‘数字鸿沟’，造成个人权利与算法权力之间的失衡，从而诱发了一系列的伦理问题。”<sup>〔112〕</sup>我们不能因为数字技术存在的问题就对其秉持悲观、消极的态度，选择一味地否定，也不能因为数字技术带来的数字红利，就允许人的主体性被数字技术消解。应当逐步摒弃“计算中立”的二分立场，选择对“人在回路”的立场进行改造，纾解“权力—权利”失衡的格局，明确在智能行政中“最终由人来面对每一位评价者，侧耳倾听他们发出的声音”<sup>〔113〕</sup>。自动化决策拒绝权作为人与数字技术对话的一次重要机会，不应人为增强二者的对抗

〔109〕 参见张欣：《算法行政的架构原理、本质特征与法治化路径：兼论〈个人信息保护法（草案）〉》，载《经贸法律评论》2021 年第 1 期。

〔110〕 参见解正山：《算法决策规制——以算法“解释权”为中心》，载《现代法学》2020 年第 1 期。

〔111〕 参见解正山：《算法决策规制——以算法“解释权”为中心》，载《现代法学》2020 年第 1 期。

〔112〕 郑智航：《人工智能算法的伦理危机与法律规制》，载《法律科学（西北政法大学学报）》2021 年第 1 期，第 14 页。

〔113〕 〔日〕山本龙彦：《机器人、AI 剥夺了人的尊严吗？》，载〔日〕弥永真生、宍户常寿编：《人工智能与法律的对话 3》，郑超、郭美蓉等译，上海人民出版社 2021 年版，第 76 页。

性，相反可以选择一种折中进路，允许人在自动化决策的任何阶段发声，捍卫人在自动化决策乃至智能行政中的尊严与主体性地位。

---

**Abstract:** With the iterative update of digital technology, the application scope and degree of automated decision-making have been continuously expanded, which has brought serious challenges to the modern administrative rule of law. Although the “right to object to automated decision-making” stipulated in Article 24, Paragraph 3 of the PIPL is relatively crude in structure and the design of the provisions is not flexible enough, it seems to provide a new direction for the study of how to protect the legitimate rights and interests of administrative counterparts in the context of intelligent administration. Embedding the right to refuse automated decision-making in intelligent administration is not only a necessary solution to the asymmetry of power and rights, an inevitable measure for embedding higher-intensity manual intervention in intelligent administration, but also a governance path of “human loop” in automated decision-making. The ban path or the right path is excluded from the scope of the right to object to automated decision-making due to different degrees of limitations. The compromise approach, with its macroscopic, weak confrontation and dynamic advantages, makes up for the limitations of the other two paths, and becomes the approach choice of the right to reject automated decision-making. Under the guidance of the compromise approach, the scope of application of the right to object to automated decision-making should be broadened, the premise of the application of “having a significant impact on personal interests” needs to be delineated, and the connection of the application of the right to object to automated decision-making should be optimized.

**Key Words:** intelligent administration, the right to object to automated decision-making, “Man in the loop”, a compromise approach, rules apply

---

(责任编辑：刘 权)