

算法解释在民法中的体系定位与类型区分

胡巧莉 刘征峰*

内容提要：从民法视角观察，个人信息应作为法益保护。针对自动化决策的算法解释实际上是人格权衍生的非独立请求权，无需单独创设一项算法解释权。算法解释在合同领域中可被纳入保护义务的范畴，在侵权领域中则具有过错评价的功能，当个人信息处理者无法解释或拒绝解释时，可推定其存在过错。为保护个人信息处理者的商业秘密，减轻其解释负担，《个人信息保护法》第24条第3款所规定的算法解释应区分为事前解释与事后解释两种类型，两者在要件构造与解释程度上存在区别。事前解释应限于“可能产生重大影响”的敏感个人信息处理场合，事后解释则发生在“对个人权益已经产生重大影响”的自动化决策场合。事前解释义务的程度应低于事后解释义务，二者分别对应系统功能的抽象解释和对具体决策的详细解释。因重大影响产生实际损害后，个人信息主体可通过主张违约或侵权损害赔偿实现救济。对事前解释义务的意定免除若违反《民法典》第497条或第506条应为无效，但事后解释义务及其衍生的损害赔偿请求权可被免除。

关键词：算法解释 个人信息权益 自动化决策 重大影响

在利用个人信息的自动化决策场合，难免在信息剖析时因“歧视性”因素而作出对个人存在重大影响或产生法律效力的决策。^{〔1〕} 规制算法的方式是将存在实质性歧视、隐私风险以及其他重大影响的决策结果纳入法律调整的范围，如个人信息主体可要求作出“算法解释”。^{〔2〕} 域外法上对“算法解释”的讨论聚焦于“算法解释（权）”的规范依据与权利性质。^{〔3〕} 欧盟以《一般数据

* 胡巧莉，中南财经政法大学法学院硕士研究生；刘征峰，中南财经政法大学法学院副教授。

〔1〕 See Bart Custers et al. ed., *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, Springer, 2012, p. 138.

〔2〕 参见郑智航、徐昭曦：《大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例》，载《比较法研究》2019年第4期。

〔3〕 See Sandra Wachter et al., Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation, 7 *International Data Privacy Law* 76, 77-81 (2017); Paul Vogel, A “Right to Explanation” for Algorithmic Decisions?, in Amedeo Santosuosso, Giulia Pinotti ed., *Data-Driven Decision Making. Law, Ethics, Robotics, Health*, Pavia University Press, 2020, pp. 49-52.

保护条例》(下文简称 GDPR)第22条和序言第71条为依据,保障个人有权不受对其产生法律效力或者类似重大影响的自动化决策之约束。为维护市场交易语境下的数据公平(或称免于受欺诈或不公平对待),美国以《信贷机会均等法》(The Equal Credit Opportunity Act)和《公平信用报告法》(Fair Credit Reporting Act)作为“算法解释(权)”的主要规范依据,专门规定了贷方需就不利的算法评分向消费者解释的义务。

我国立法文本中未出现“算法解释(权)”这样的表达,但根据《个人信息保护法》第24条第3款的规定,个人信息主体有权要求个人信息处理者说明对其权益有重大影响的自动化决策,并享有相应的拒绝权。值得探讨的是,该款规定隐含的“算法解释”在民法上究竟属于何种性质,又可产生何种法律效果。我国学界对“算法解释”是否为一项独立的民事权利存在争论,肯定者认为个体赋权式的治理模式可为个人信息提供更强保护,^[4]并由此形成完整的个人信息权利体系。否定者则认为个人信息仅需作为利益保护,算法解释的功能可被现有的法律制度涵盖,创设算法解释权易形成保护功能的堆叠。^[5]两种相反的观点虽各有理由,但未能就民法体系中如何理解算法解释作出圆满的回答。无论采何种规范模式,核心是如何在主观权利体系中厘清“算法解释”的保护目的与体系价值。

本文以利用个人信息的自动化决策为研究对象,首先从“算法解释”保护的个人信息权益的法律评价出发,判断算法解释能否被人格权中的个人信息保护(《民法典》第1032条至第1039条)、消费者知情权(《消费者权益保护法》第8条、第14条以及《个人信息保护法》第44条)、合同保护义务、侵权责任等民法保护模式的规范功能所涵括,明晰其性质与定位。在此基础上,尝试在民法体系中对算法解释进行类型区分,重点分析类型区分的实益以及不同类型算法解释的规范构造。

一、人格权体系中的算法解释

《个人信息保护法》第24条规定了利用个人信息的自动化决策和相应的“算法解释”义务。算法解释作为保护个人信息权益的方式,其体系定位与规范性质尚未明确,关键在于目前的民法体系中是否已经存在可以承载“算法解释”的规范制度。因此,首先,需要先明确个人信息权益的性质,再判断算法解释的体系位置是处于“个人信息保护体系”还是人格权保护体系,其性质为独立权利还是请求权。其次,基于算法解释的规范目的,自动化决策对个人权益的重大影响产生的时间不同,需区分不同的算法解释类型。

(一) 作为非独立请求权的算法解释

《民法典》规定自然人的个人信息受法律保护,列明了个人信息处理中需遵循的原则和违反相应义务时需承担的责任。《个人信息保护法》对个人信息保护作出了更为详尽的规定,但均未

[4] 参见张凌寒:《商业自动化决策的算法解释权研究》,载《法律科学(西北政法大學學報)》2018年第3期;张恩典:《大数据时代的算法解释权:背景、逻辑与构造》,载《法学论坛》2019年第4期;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。

[5] 参见贾章范:《论算法解释权不是一项法律权利——兼评〈个人信息保护法(草案)〉第二十五条》,载《电子知识产权》2020年第12期;辛巧巧:《算法解释权质疑》,载《求是学刊》2021年第3期。

言明“个人信息权”的存在，仅出现“个人信息权益”的表达。个人信息属于人格要素，与人格权的保护内容存在交叉，若重复性地创设独立的“个人信息权”，则会与人格权的保护体系产生叠合。^{〔6〕}立法政策对个人信息保护采行为规制模式，正是由于个人信息权益尚未达致“权利的分配密度”（Zuweisungs-dichte）。^{〔7〕}行为规制模式将个人信息视为一项法益，规制信息处理行为，既可避免权利泛化，亦可缓解与个人信息利用与保护之间的紧张关系。^{〔8〕}因此，个人信息法益可置于人格权规范体系下予以保护，尽管《个人信息保护法》第四章“个人在个人信息处理活动中的权利”中使用了“权利”，但其仅为技术意义上的表达，与《民法典》第五章“民事权利”中具体列举的生命权、身体权等权利的性质并不相同，将其理解为保护个人信息法益的请求权更为适宜。

在明确了个人信息权益属于人格权保护体系的法益后，作为其保护方式的“算法解释”，亦应当被置于人格权保护体系的范畴进行性质界定。对于“算法解释”的关键争议在于《个人信息保护法》第24条第3款是否创设了“算法解释权”。支持的观点认为目前的法律资源不敷适用，应配置独立的算法解释权，用以平衡自动化决策使用者与相对人不对称的权力关系。^{〔9〕}只有通过赋权才能增强算法透明性，以破解“算法黑箱”的方式达到提高算法可责性的目的。^{〔10〕}否定的观点认为算法解释的功能可通过其他规制手段实现，^{〔11〕}创设权利的救济效果有限，会对商业秘密、国家利益和市场竞争秩序造成不良影响，徒增裁判成本。^{〔12〕}深思之，支持算法解释“权利化”的观点未从民法体系的视角去观察算法解释的定位。算法解释所保护的法益可以被人格权规范体系所涵盖，在合同与侵权领域表现为不同的保护手段与规范功能。至于商业秘密，其仅构成对算法解释说明义务的限制，不能绝对地拒绝算法解释请求权的适用。^{〔13〕}

域外法上，学者们以GDPR第22条及序言第71条为核心对算法解释的规范模式展开了讨论。有观点认为即便序言在欧盟法律下并不具有拘束力，但为了引导数据主体依据GDPR行使权利，至少应保障其从功能角度获得解释的利益。^{〔14〕}相反观点则认为GDPR第22条并未明文规定“算法解释权”，序言第71条虽提及“算法解释权”（the right to obtain an explanation of the decision），但其未处于正式文本中，不可创设权利类型。^{〔15〕}亦有观点主张“算法解释权”可能引发新的“透明性谬误”，通过被遗忘权、数据可携带权、隐私保护等即可使算法更加注重

〔6〕 参见郑晓剑：《个人信息的民法定位及保护模式》，载《法学》2021年第3期；程啸：《论我国民法典中个人信息权益的性质》，载《政治与法律》2020年第8期。

〔7〕 Vgl. Schwab/Löhnig, Einführung in das Zivilrecht, 20. Aufl., 2016, Rdn. 263.

〔8〕 参见黄薇主编：《中华人民共和国民法典人格权编释义》，法律出版社2020年版，第196页。

〔9〕 参见前引〔4〕，张凌寒文；前引〔4〕，解正山文。

〔10〕 参见姜野、李拥军：《破解算法黑箱：算法解释权的功能证成与适用路径——以社会信用体系建设为场景》，载《福建师范大学学报（哲学社会科学版）》2019年第4期；许可、朱悦：《算法解释权：科技与法律的双重视角》，载《苏州大学学报（哲学社会科学版）》2020年第2期。

〔11〕 参见前引〔5〕，辛巧巧文。

〔12〕 参见前引〔5〕，贾章范文；李天助：《算法解释权检视——对属性、构造及本土化的再思》，载《贵州师范大学学报（社会科学版）》2021年第5期。

〔13〕 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。

〔14〕 See Bryan Casey et al., Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise, 34 Berkeley Technology Law 143, 159-161 (2019).

〔15〕 参见前引〔3〕，Sandra Wachter等文，第79-80页。

个体利益保护。^[16] 不同成员国对算法解释的规范模式并未与 GDPR 完全保持一致。德国、比利时并未肯认 GDPR 序言第 71 条所提及的“算法解释权”。^[17] 可能的原因在于,基于算法自动化决策所作出的歧视性结果可由其他法律规范进行规制。如德国《一般平等待遇法》不仅将损害赔偿作为救济方式,而且将救济方式扩展至可能对契约自由作根本性否定的强制缔约。^[18] “如果歧视行为表现为拒绝订立合同,在只有通过订立合同才能排除妨碍的情况下,排除妨碍的外延中实际上包含了强制缔约这一方式。”^[19] 在英国与爱尔兰的《数据保护法》中,虽不能直接看出“算法解释权”,但其均要求数据控制者阐明算法分析的过程。^[20] 法国和匈牙利的相关法令将“算法解释权”纳入了规范范畴,只是在行使的具体要求上略有不同。^[21] 甚者,意大利最高上诉法院在判决中直接将“算法是否透明”作为个人信息处理中的“个人同意”是否有效的关键因素。若算法不透明则个人信息主体作出的“同意”归于无效,对利用算法进行自动化决策的手段进行严格规制。^[22]

还有学者倾向于将概念限缩,不称其为“解释权”(right to explanation)而称其为“知情权”(right to be informed),认为知情内容包括算法系统功能与具体的自动化决策。^[23] 但知情权对应的是个人信息处理者的事前告知义务,披露的时间为收集信息时、处理个人信息前,不包括处理个人信息并作出决策后的情形。^[24] 在内容上,知情权旨在披露个人信息处理的主体、目的、内容、方式以及享有的访问权等权利,^[25] 不包括对具体自动化决策结果的解释。我国《个人信息保护法》第 44 条规定个人对其个人信息的处理享有知情权,并未言明知情的具体内容。第 17 条对告知义务内容的规定亦不包含针对个人的具体自动化决策。

因此,算法解释并不等同于算法透明原则项下的知情权。知情权的内容较算法解释更为狭窄,不要求“重大影响”作为要件,亦不包含复杂的事后解释类型。若不对事前与事后进行区分,统一以知情权保护,将会忽略对个人信息处理者商业秘密的保护和解释负担的关注。在规范效果上,算法解释不仅具备与告知义务类似的事前防御性效果,还有事后救济的功能,具体表现为合同领域中义务不履行责任的判断和侵权责任领域中过错的评价。无论是《消费者权益保护法》中的消费者知情权,还是《个人信息保护法》中的个人信息主体的知情权,均无法完全囊括

[16] 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 & Technology Review* 18, 19–20 (2017–2018).

[17] See Gianclaudio Malgieri, Automated Decision-making in the EU Member States: The Right to Explanation and other “Suitable Safeguards” in the National Legislations, 35 *Computer Law & Security Review* 1, 8–13 (2019).

[18] 参见刘征峰:《从“反歧视原则”进入民事交易关系观察当代民法理念的革新》,载《法制与社会发展》2017年第1期。

[19] Haberl, Antidiskriminierungsrecht und Sanktionensystem: Die Konkretisierung gemeinschaftsrechtlicher Mindestvorgaben, GPR 6 (2009), 202, 202–209.

[20] See UK Data Protection Act 2018, Chapter 2, Section 14; Irish Data Protection Act 2018, Part 3, Chapter 3, Article 57.

[21] 参见前引 [17], Gianclaudio Malgieri 文,第 21–22 页。

[22] See Giorgia Bincoletto, Italy Supreme Court of Cassation on Automated Decision Making: Invalid Consent if an Algorithm is Not Transparent, 7 *European Data Protection Law Review* 248, 248–249 (2021).

[23] 参见前引 [14], Bryan Casey 等文,第 161 页。

[24] See Isak Mendoza, Lee A. Bygrave, The Right Not to Be Subject to Automated Decisions Based on Profiling, in Synodinou et al. ed., *EU Internet Law: Regulation and Enforcement*, Springer, 2017, p. 93.

[25] See GDPR Article 13, Article 14; 我国《个人信息保护法》第 17 条。

算法解释的规范功能。

综上所述,判断“算法解释(权)”是否存在,不应拘泥于立法文本的形式表达,应当充分结合背后的规范目的和价值考量。从《民法典》所采的立场来看,作为一般人格权条款的第990条第2款以“其他人格权益”保护一般人格权,与该条第1款所规定的具体人格权相区分。^[26]个人信息保护的诸多内涵可被纳入具体人格权与一般人格权益的保护范畴。因此,在不存在独立的个人信息权的前提下,算法解释作为保护个人信息权益的方式,既不是作为个人信息权的权能,也不是作为独立的算法解释权,其民法性质为保护个人信息法益的非独立请求权。^[27]

进一步而言,算法解释是依附于人格权保护体系、为实现人格利益保护而服务的非独立请求权。《个人信息保护法》第24条第3款和第48条规定的解释义务以及第44条所规定的“知情权”存在保护内容上的重合之处,均构成个人信息主体在个人信息处理中要求事前知情或者事后解释的权益保护手段。对于《个人信息保护法》与《民法典》之间的适用关系,合理的模式应为《民法典》人格权编仅对个人信息保护作出原则性和转介性的规定,《个人信息保护法》对具体问题作出更为细致的规范。在《个人信息保护法》中没有明文规定时,才可适用《民法典》中的一般条款。^[28]

(二) 非独立请求权性质下的类型区分实益

正如前文所言,若不对算法解释进行事前事后的区分,可能导致其与知情权的混淆和对事后救济功能的忽视。个人信息属于人格要素,而保护人格权的方式在事前和事后有所区别,损害产生的事前表现为防御型请求权,损害产生的事后表现为救济型请求权。因此,值得探讨的是:算法解释作为人格权保护项下的非独立请求权,是否应参照人格权保护方式因事前事后的差异而区分不同的类型。

GDPR未对算法解释的性质属于事前解释还是事后解释作出明确规定。^[29]事前解释发生于算法提供者基于用户同意开始收集个人信息后、自动化决策作出前,决策系统的规则设计可能对个人信息主体权益产生重大影响。事前解释提供的信息为系统功能,例如可能被考虑的数据类型及特征或决策树的分级标准等,未具体到个人信息主体。^[30]事后解释则发生在自动化决策已经作出,且对个人信息主体的个人权益产生了重大影响的情形。事后解释的内容既包括系统功能也包括针对个人的具体决策,如某一特定主体被采纳的信息类型、个案中纳入自动化决策的因素和权重、对权益产生的重大影响等。^[31]

与合同法上的保护义务以及侵权责任领域的交往安全义务相类似,只有当个人信息主体与个人信息处理者进入到一定程度的特别结合关系中,才可对其施加解释义务的负担,不可在尚未开

[26] 参见朱晓峰:《论一般人格权条款与具体人格权条款的规范适用关系》,载《比较法研究》2021年第3期。

[27] “请求权有独立请求权与非独立请求权之分,独立的请求权不依赖于其他权利而独自存在,本身即属于一种权利,而非独立请求权则是为实现其他的权利服务的。”参见〔德〕卡尔·拉伦茨:《德国民法通论》(上册),王晓晔等译,法律出版社2013年版,第325页。

[28] 参见王苑:《个人信息保护在民法中的表达——兼论民法与个人信息保护法之关系》,载《华东政法大学学报》2021年第2期;石佳友:《个人信息保护的私法维度——兼论〈民法典〉与〈个人信息保护法〉的关系》,载《比较法研究》2021年第5期。

[29] 参见前引〔3〕,Sandra Wachter等文,第82页。

[30] 参见前引〔3〕,Sandra Wachter等文,第78页。

[31] 参见前引〔17〕,Gianclaudio Malgieri文,第22-23页。

始收集个人信息时即要求作出事前解释。但我国《个人信息保护法》第24条第3款忽略了事前解释与事后解释在保护目的和解释内容上的区别。本文认为,应作事前解释和事后解释的类型区分,区分的实益如下:

首先,事前解释的功能无法被《个人信息保护法》第17条的事前告知义务所涵盖,该条第1款第2项规定,个人信息处理者在处理个人信息前,应当以显著方式、清晰易懂的语言真实、准确、完整地向个人信息主体告知个人信息的处理目的、处理方式、处理的个人信息种类、保存期限。其中的处理方式仅意味着选择人为还是自动化决策,^[32]并不要求具体解释算法系统功能以及算法运行规则。《个人信息保护法》第48条规定个人信息主体可要求个人信息处理者解释说明的“个人信息处理规则”,与第17条的“处理的方式”亦不相同。

有学者提出“算法说明义务”这一概念,认为从告知义务中可推导出“算法说明义务”,其适用于算法决策的全过程,以避免陷入算法解释是否为独立权利类型的争议。算法解释仅包括自动化决策已经对个人权益产生重大影响的事后解释,而“算法说明义务”则不拘泥于“重大影响”这一条件。^[33]该观点存在以下可供商榷之处:第一,“算法说明义务”是一个较为模糊的概念,其简单地将“告知义务”与“算法说明义务”等同,但无论是《个人信息保护法》第7条、第14条还是第17条均无法解释“算法说明义务”的具体内涵。第二,通过模糊概念规避算法解释的性质分析存在不妥。如前文所述,算法解释的性质应为人格权保护体系中的非独立请求权。《个人信息保护法》第24条第3款不仅包括事后解释,亦包括自动化决策作出前的事前解释。事前解释与事后解释在适用要件上有所区别,前者仅要求“可能的重大影响”,后者要求“已经产生重大影响”。第三,前文已述,《个人信息保护法》第17条规定的告知义务仅为事前告知义务,若将“告知义务”与“算法说明义务”等同,则“算法说明义务”无法包含事后说明的类型。第四,若“算法说明义务”无要求“重大影响”这一要件,会过度增加个人信息处理者的解释负担,不利于个人信息的流通。因此,通过“算法说明义务”无法否定算法解释存在事前解释和事后解释两种类型。

其次,事前解释注重防御性的保护,目的在于避免抽象的重大影响转化为针对个人的重大影响;事后解释则注重对已经发生的重大影响作出解释,并在重大影响转化为损害后对义务违反及过错要件进行评价。事前解释作为防御性请求权,可以为损害实际产生前的个人信息主体提供预防性的保护,尤其是以敏感个人信息为对象的处理行为。这与《民法典》第997条规定的防御性请求权的规范功能相似。该条的根本目的在于为人格权保护提供高效、便捷的预防性保护措施。^[34]对于自动化决策中个人权益的保护亦应如此。虽然此时损害尚未发生,但损害赔偿带来的负担不仅仅发生在具体加害人与受害人之间,还会造成额外的社会支出,预防损害能够以最低成本产生最大收益。^[35]不过,事前解释的适用场景需作限缩,否则会对个人信息处理者造成过

[32] 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179页。

[33] 参见前引[13],吕炳斌文。

[34] 参见程啸:《论我国民法典中的人格权禁令制度》,载《比较法研究》2021年第3期。

[35] 参见朱岩:《侵权责任法通论·总论》,法律出版社2011年版,第107页;〔奥〕海尔姆特·库齐奥:《侵权责任法的基本问题(第一卷):德语国家的视角》,朱岩译,北京大学出版社2017年版,第78-80页。

重的解释负担。

再者，区分事前解释与事后解释更有利于实现对商业秘密的保护。在自动化决策作出前，个人权益尚未受到重大影响，应基于比例原则承认商业秘密对解释请求权的一般性抗辩。“合法、正当、必要原则”已成为个人信息处理的基本原则（《个人信息保护法》第5条），是比例原则在民法中的体现，有益于弥补数字时代意思自治与契约自由的缺陷。^{〔36〕}比例原则的适用不应局限于个人信息处理者的处理行为，还应包括对个人信息主体算法解释请求权的限制。例如，作为框架权的一般人格权的效力弱于具体人格权，只有通过利益权衡，才能认定是否存在侵害这些权益的行为。在进行利益权衡时，必须考虑到其他人与此相冲突的权益。^{〔37〕}基于人格权所衍生的非独立请求权更应如此，当事前解释与他人的商业秘密产生冲突时，由于此时对个人信息主体权益的重大影响尚未实际产生，故个人信息处理者的商业秘密应优先受到保护。

正如GDPR序言第63条所指明的，个人信息权利的行使不应对其他主体的合法权益产生不利影响，尤其是商业秘密、知识产权以及软件版权。^{〔38〕}实践中，多数个人信息处理者在拒绝作出解释时给出的理由为“保护商业秘密”。如在“卢米斯案”中，上诉法院因“商业秘密保护”而拒绝了被告所提出的“打开再犯风险评估算法”的请求。^{〔39〕}在“SCHUFA案”中，当事人因较低的自动化信贷评分而失去了借贷机会，但其在申请评估公司公开计算评分的算法规则时，同样被法院以保护商业秘密的理由而驳回。^{〔40〕}算法本身具有客观性，且往往蕴含着商业秘密，这已成为增强算法问责制和透明度的主要障碍之一。^{〔41〕}有观点认为，只有在算法解释涉及公共利益的情形下才可考虑进行公开。^{〔42〕}

我国《个人信息保护法》第48条未对解释说明的限度、个人信息处理者的抗辩事由作出规范。在实践中，如果缺乏法律的明确规定，平衡算法解释与商业秘密将会面临较大的困境。^{〔43〕}商业秘密保护有效地在算法中创造了一种不需要公开的财产权益，但这并非意味着只要进行算法解释就会不利于商业秘密的保护。首先，算法解释的内容并非均属于商业秘密的范畴。复杂的代码与运算模型对于非专业人员而言并无解释的实质效果，只有将算法解释与具体的决策结果相联系，才能够达到个人信息主体主张解释的目的。其次，算法解释所寻求的是“有法律意义的信息”，通过对解释对象和解释内容的限制可以规避侵害商业秘密的风险。在解释对象上，可针对某些特定群体进行小范围的公开解释。在解释内容上，无需对核心算法进行公开解释，但至少应当公开影响算法决策过程的决定性因素及一般操作模式。^{〔44〕}在解释条件上，限于法律明确规定

〔36〕 参见刘权：《论个人信息处理的合法、正当、必要原则》，载《法学家》2021年第5期。

〔37〕 参见〔德〕迪特尔·梅迪库斯：《德国民法总论》，邵建东译，法律出版社2013年版，第107页。

〔38〕 See GDPR Recital 63 sentence 5.

〔39〕 See *State v. Loomis*, 881 N. W. 2d 749 (Wis. 2016).

〔40〕 Vgl. BGHZ 200, 38.

〔41〕 See Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 *U. C. L. A. Law Review* 54, 54–59 (2019).

〔42〕 See *Campbell v Frisbee* [2002] EMLR 31; Guido Noto La Diega, *Against the Dehumanisation of Decision-Making: Algorithmic Decisions at The Crossroads of Intellectual Property, Data Protection, and Freedom of Information*, 9 *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 3, 13 (2018).

〔43〕 See Gianclaudio Malgieri, *Trade Secrets v Personal Data: A Possible Solution for Balancing Rights*, 6 *International Data Privacy Law* 102, 102–106 (2016); 前引〔42〕, Guido Noto La Diega文，第14–18页。

〔44〕 参见前引〔3〕, Paul Vogel文，第54–55页。

的情形或者仅在个案中向单个数据主体披露。^{〔45〕}因此,将可能涉及商业秘密的解释内容限制在事后解释的范畴则可避免对个人信息处理者商业秘密的过度妨害。

综上所述,应以对个人权益的重大影响或损害实际发生的时间为节点,对算法解释的类型作出区分,分为收集信息后决策作出前可能产生重大影响的事前解释与决策作出后并产生重大影响的事后解释两种类型。类似于合同缔结前后双方当事人所承担的保护义务内容与程度的差异,事前解释的义务低于事后解释,以平衡个人信息主体的合法权益与个人信息处理者的商业秘密保护等利益。

二、事前的算法解释与基于同意的义务免除

根据《个人信息保护法》第24条第3款的规定,算法解释需遵循的规范要件有二:第一,自动化决策。《个人信息保护法》第24条第3款中“拒绝”所指向的是“仅通过自动化决策的方式作出决定”,不包括掺杂人工干预的自动化决策结果。第二,该决策对个人权益有重大影响。个人权益并不限于个人信息权益,亦包括《民法典》总则编第5章所规定的其他民事权利,如生命权、身体权、健康权、财产权等。因此,在事前解释与事后解释中,自动化决策方式与个人权益是两者的共同要件,但两者对“重大影响”的判断并不相同。

(一) 敏感个人信息处理中对重大影响的推定

事前解释中的“重大影响”为“可能的重大影响”。对于“重大影响”这一要件,《个人信息保护法》第24条第3款与GDPR第22条的规范表述并不相同。GDPR第22条将数据主体不受自动化决策限制的要件列为“对数据主体产生法律效力或对其造成类似的重大影响”,从文义结构上看,产生法律效力应属于“重大影响”的一种具体表现。我国《个人信息保护法》第24条第3款仅要求“重大影响”,但从规范目的而言,“重大影响”应包含产生法律效力这一种表现形式。

遍阅《民法典》,并无“重大影响”这样的表达,与此类似的仅有《民法典》第496条中的“重大利害关系”。“重大利害关系”的认定可参考《消费者权益保护法》第26条第1款,如商品或者服务的数量和质量、价款或者费用、履行期限和方式、安全注意事项和风险提示、民事责任等内容属于与消费者有重大利害关系的范畴。通常认为,《民法典》第470条规定的合同一般条款均属于与当事人有重大利害关系的条款。^{〔46〕}具体到个人信息保护领域,若个人信息处理者基于同意而实施个人信息处理行为,“重大利害关系”不仅表现为同意范围内的事项,而且包括可能对个人信息主体的民事权利产生重大影响的内容。若个人信息处理者基于订立合同必需而实施信息处理行为,“重大利害关系”则主要表现为合同这一基础关系所构建的权利义务关系。抽象而言,“重大利害关系”涵括任何对个人信息主体民事权益产生影响的内容。这样的理解也进一步回应了“重大影响”应该包含“产生法律效力”的解释。

《个人信息保护法》第55条及第56条对于个人信息保护影响评估作出了规范,要求个人信

〔45〕 参见前引〔4〕,解正山文。

〔46〕 参见最高人民法院民法典贯彻实施工作领导小组主编:《中华人民共和国民法典合同编理解与适用》〔一〕,人民法院出版社2020年版,第246页。

息处理者在事前应当进行个人信息保护影响评估，评估的内容包含“对个人权益的影响及安全风险”。具体而言，《信息安全技术个人信息安全影响评估指南》（GB/T 39335—2020）第 5.5.1 对个人权益维度进行了细化。对个人权益的影响可分为以下四类：第一，限制个人自主决定权，例如无法选择拒绝个性化广告的推送，被蓄意推送影响个人价值观判断的资讯等。第二，引发差别性待遇，例如因疾病、婚史、学籍等信息泄露造成的针对个人权利的歧视。第三，使个人名誉受损或遭受精神压力，例如被他人冒用身份、监视追踪等。第四，人身财产受损，例如引发人身伤害、遭受诈骗、勒索等。这些可能对个人权益产生重大影响的内容与《个人信息保护法》第 28 条中的“人格尊严受到侵害、人身财产安全受到危害”高度吻合。因此，在事前解释中，通过个人信息的敏感度即可推定对个人权益的重大影响。

从《个人信息保护法》第 28 条至第 30 条的规定可以明显看出法律对敏感个人信息的特殊保护。敏感个人信息是一旦泄露或者非法使用则容易导致自然人的人格尊严受到侵害或者人身、财产安全受到危害的个人信息（《个人信息保护法》第 28 条），在个人信息处理过程中容易对个人权益产生重大影响。基于此，个人信息处理者处理敏感个人信息的，应当向个人告知处理敏感个人信息的必要性以及对个人权益的影响，但在一般个人信息的处理中，并未要求个人信息处理者事前告知自动化决策对个人权益的影响。损害尚未实际发生时，对个人信息主体的保护力度应低于损害实际发生的时候。因此，需通过区分个人信息的类型对事前解释作出限制，事前解释的对象应限于对敏感个人信息的处理行为。非敏感的个人信息一般不会引起对个人权益的重大影响，且在事前难以证明。若将非敏感的个人信息处理纳入事前解释的范畴，会对个人信息处理者课加过重的解释负担。

在符合上述要件后，事前解释的具体内容不仅包括《个人信息保护法》第 17 条的内容，亦应包括向个人信息主体披露自动化决策适用的逻辑和可能的一般性后果。^[47]但个人信息处理者仅需作抽象解释，并不具体到与相对人有关的自动化决策结果。^[48]欧洲法院对算法解释的内容采取了较为谨慎的态度，其指出责任主体仅需对一般性的信息以可理解的方式作出解释。^[49]事前解释的目的在于帮助个人信息主体了解预期或未来的自动化决策，以便在知情的基础上决定是否允许自己的个人信息被处理、评估自动化处理的合理性或者行使更正权、反对权等请求权。^[50]通过对系统功能等抽象内容的解释，已经足够达到相应的防御效果。^[51]

（二）个人同意与解释义务的免除

若个人信息主体事前同意免除个人信息处理者的算法解释义务，会对该义务产生何种影响？免除算法解释义务的事前同意与《个人信息保护法》第 13 条第 1 款第 1 项中规定的“取得个人

[47] See GDPR Article 13 (2) (f)、Article 14 (2) (g).

[48] 参见前引 [10]，姜野、李拥军文。

[49] See *YS v Minister voor Immigratie, Integratie en Asiel* (C-141/12) and *Minister voor Immigratie, Integratie en Asiel v M, S* (C-372/12), ECJ judgment of 17/07/2014 para. 50 et seq.

[50] See Sandra Wachter et al., Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR, 31 *Harvard Journal of Law & Technology* 841, 865-867 (2018).

[51] See Antoni Roig, Safeguards for the Right not to be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR), 8 *European Journal of Law and Technology* 1, 3 (2017).

的同意”并非等同的概念。前者同意的内容为免除将来的算法解释义务，是对责任的免除；而后者仅同意个人信息处理者对个人信息的处理行为，并非对责任的免除。对于何种情形下能够免除个人信息处理者的解释义务，目前并无细致的研究，仅有学者就免除告知义务主张应当以个人信息是否直接向信息主体收集作为区分标准。^{〔52〕}但这与解释义务的免除并不相同，事前告知义务的内容仅可被事前算法解释所涵括，忽视了事后解释的类型。无论是告知义务的免除还是解释义务的免除，均属于对个人信息主体权益的处分，在适用情形上可参考类似的规范。

《个人信息保护法》第24条第3款未对算法解释请求权的行使设置例外。GDPR第22条第2款设置了3项例外情形，第一项是“该决定为订立、履行个人作为一方当事人的合同所必需”，合同作为双方当事人分配风险的工具，若该自动化决策系双方经过自由磋商而订入合同且合同不存在无效情形，则可视为对抗解释义务的事由。第二项是“该决定是由欧盟或其成员国的法律授权作出的，且该欧盟或其成员国的法律不仅约束控制者，并且明确了保护数据主体权利、自由和合法利益的适当措施由联盟或成员国的法律规定”，此种例外情形是指存在明确的法律授权保障该自动化决策的实施，置于我国法律规范的语境下类似于“法律、行政法规规定的其他情形”。如果该自动化决策的决定系为履行法定职责或者法定义务所必需，除非该法定职责或者义务包含无需针对自动化决策作出解释的内容，一般不允许其直接对抗解释义务。但我国《个人信息保护法》第13条第1款第4、5项的内容系基于保护公共利益的目的而设置，公共利益的保护可通过限制个人信息处理者的信息处理行为而实现，限制个人信息处理者的算法解释义务并不能达到此种规范目的，因此无需将其作为免除解释义务的场合处理。第三项是“该决定基于个人信息主体的明确同意”，此处的“明确同意”与个人信息保护中的“告知—同意”规则并不相同。“告知—同意”规则仅针对《个人信息保护法》第17条的告知事项，并不包含解释义务中所指向的算法规则与具体决策结果之间的关联。“明确同意”的实质为个人信息主体明知自动化所作出的决定可能会或者已经对个人权益产生重大影响仍愿意接受该种结果的发生。

关键在于，个人信息主体通过事前同意免除解释义务是否有效。个人同意系属于私主体对自己个人信息权益的处分，^{〔53〕}对于个人信息保护中个人信息主体的“同意”的性质，学界存在诸多争议，大致为“法律行为许可说”^{〔54〕}“信托授权行为说”^{〔55〕}“持续性代理行为说”^{〔56〕}等。从民法的角度而言，“同意”的性质应当区分合同与侵权不同的场合而定。在合同领域，同意可能成为合同中给付内容的一部分。在侵权领域，同意或可作为免责事由（受害人同意）免除个人信息处理主体的责任（发生在事后解释的情形下）。^{〔57〕}

通过事前同意免除解释义务的方式分为三种，其一为个人信息处理者提供格式条款排除其解

〔52〕 参见程啸：《论个人信息处理者的告知义务》，载《上海政法学院学报（法治论丛）》2021年第5期。

〔53〕 参见万方：《个人信息处理中的“同意”与“同意撤回”》，载《中国法学》2021年第1期。

〔54〕 参见刘召成：《人格商业化利用权的教义学构造》，载《清华法学》2014年第3期。

〔55〕 参见丁晓东：《个人信息权利的反思与重塑——论个人信息保护的适用、前提与法益基础》，载《中外法学》2020年第2期。

〔56〕 See Jennifer Barrigar et al., Let's Not Get Psyched Out of Privacy: Reflections on Withdrawing Consent to the Collection, Use and Disclosure of Personal Information, 44 *Canadian Business Law Journal* 54, 60 (2006).

〔57〕 参见前引〔53〕，万方文。

释义务，其二为个人信息主体与个人信息处理者磋商达成免除解释义务的约定，其三为个人信息主体单方免除个人信息处理者的解释义务。第一种方式较为常见，典型情形为网络平台设置格式条款，如“用户若同意上述内容，则本平台无需对自动化决策所作出的决定承担解释责任”。个人信息主体仅有是否接受免除解释义务的选择空间，并无磋商余地。但格式化条款可能因排除个人信息主体的主要权利而无效。根据网络平台对自身解释义务减轻或免除的不同程度，分别可能符合《民法典》第497条第2项或第3项，即个人信息处理者通过格式条款不合理地免除或者减轻其责任、加重对方责任、限制对方主要权利和信息处理者直接排除对方主要权利的情形。个人信息主体对个人信息处理者所享有的算法解释请求权是否属于“主要权利”？前文虽否定独立的算法解释权，但算法解释在合同领域中可辅助主权利的实现，对该义务的违反会造成利用个人信息提供服务这一主给付义务的不适当履行。因此，对算法解释的保护应纳入该条所规定的“主要权利”的范畴。第二种方式是个人信息处理者与个人信息主体通过磋商达成免除解释义务的约定。解释义务产生的前提是对个人权益的重大影响，若此种影响已经足以造成对方人身损害或者属于因故意或重大过失造成对方财产损失的，则因违反《民法典》第506条而归于无效。第三种方式是个人信息主体直接单方免除其解释义务，其无效的情形与第二种方式相同。

事前同意免除解释义务，是否需要严格限定于“明确同意”的情形？若个人信息主体默示同意，是否发挥相应的免除效果？在意思表示的概念中，应区分两个方面：作为行动的意思表示以及作为客观逻辑的意义构造的意思表示。只有某一举动本身是被意愿的，才存在一项民法意义上的具有法律重要性的“行动”。^{〔58〕} 设置为明确同意的意义在于，自动化决策可能对个人权益产生重大影响，若被免于解释，个人信息主体的同意实际上是对相关合同责任或侵权责任的豁免，意思表示必须清楚且确定，以避免引起权利义务关系的争议。对于默示的意思表示，由于在个人信息处理中存在个人信息主体知情权与个人信息处理者告知义务的断层，同意的意思表示本就存在不自由的空间，^{〔59〕} 若允许默示同意，则难以避免个人信息主体作出含有效力瑕疵的意思表示。

因此，个人信息处理者的解释义务不可轻易被事前免除，若自动化决策可能对个人信息主体的权益产生重大影响，个人信息处理者需对自动化决策的相关内容作出清晰、通俗的解释。《个人信息保护法》第24条第3款并未规定违反该事前解释义务的民事责任，需具体分析：若双方将解释义务订入合同约定，则依照是否违反合同义务界定民事责任；在无合同约定的场合，个人信息主体可拒绝个人信息处理者仅依据自动化决策作出决定。而事后所产生的解释义务以及损害赔偿义务可被免除，原因在于当事人可对自己已经产生的算法解释请求权及损害赔偿请求权予以处分，与事前免除的效果存在本质区别。

三、事后解释在救济体系中的功能与损害赔偿

事后解释中的关键要件是“已经对个人权益产生实际的重大影响”，且无论是一般个人信息

〔58〕 参见〔德〕卡尔·拉伦茨：《法律行为解释之方法——兼论意思表示理论》，范雪飞、吴训祥译，法律出版社2018年版，第36-37页。

〔59〕 参见前引〔53〕，万方文；武腾：《最小必要原则在平台处理个人信息实践中的适用》，载《法学研究》2021年第6期。

还是敏感个人信息，只要符合该要件即可要求事后解释。原因在于，重大影响或者损害已经实际产生，既不会对个人信息处理者造成过重的解释负担，亦有正当理由反驳以商业秘密为抗辩。至于何种结果构成“重大影响”，需在合同或侵权等不同场域分别判断。在重大影响已经实际产生的情况下，事后解释并不能发挥完全的救济效果，需辅之损害赔偿责任。

（一）事后解释中的重大影响

事后解释中重大影响的典型一般为不当处理个人信息，导致个人在工作机会、金融安排等方面遭受负面影响。^{〔60〕}一方面包括自动化决策对个人信息主体的法律地位或合同权利等产生影响，如合同撤销、社会福利丧失、被拒绝给予公民身份等；另一方面也包括个人的法律权利或义务未因自动化决策而改变，但是该决策对个人的境遇、行为或选择产生显著影响，或导致其因歧视被排除在某种机会之外（如金融服务、就业或者教育机会丧失）。^{〔61〕}但根据《个人信息保护法》第24条第3款的文义，“重大影响”并未限定于不利影响的层面，即便是产生有利的重大影响，个人信息主体仍有选择是否接受的自由。

鉴于上述讨论，何种行为可被评价为“对个人权益已经产生重大影响”乃问题的核心。个人权益的认定与影响是否重大存在直接的关联，该问题可转化为如何认定构成重大影响的个人权益。在合同关系中，个人信息处理对个人权益产生重大影响的典型表现为个人信息处理的格式条款，网络平台多基于此声明有权对个人信息进行收集、存储、使用等。因此，对《民法典》第497条中“主要权利”的判断可资借鉴，但何为“主要权利”，需要根据合同性质本身确定，不仅仅关注双方当事人签订合同的内容，而应就合同本身的性质来考察。^{〔62〕}“主要权利”不可简单地解释为法律规定的权利，或由法院基于公平原则进行自由裁量所认定的权利。^{〔63〕}因此，“对个人权益已经产生重大影响”在合同领域表现为违反主给付义务以及违反其他给付义务导致个人权益遭受损害，在侵权领域表现为对个人信息主体主要权利造成损害。

由于此时影响已经实际产生，对于影响程度的判断相较于事前解释中对重大影响的推定更为容易且准确。依据《信息安全技术个人信息安全影响评估指南》（GB/T 39335—2020）中所提供的个人权益影响程度判定准则，可以个人信息主体所遭受的困扰程度、付出的成本、生理疾病、财产权益损害、信用名誉损害等作为标准判断影响的严重程度。若该自动化决策行为使得个人信息主体遭受诈骗、信用评分和名誉受损、失去工作机会、产生生理疾病等，则可被评价为对个人权益产生了重大影响。当然，这些情况在限制个人自主决定权、引发差别性待遇、人身财产受损等不同场景中存在不同的表现，需具体判断。若只是对个人信息主体造成了精神上的厌烦与恼怒情绪，则仅被评价为低程度的影响，不符合要求事后解释的要件。

在事后解释中，算法控制者应当向个人信息主体披露的信息不限于算法系统功能的一般性事前解释，而应对具体的算法决策结果予以说明，如基础性数据、算法规则、决策结果三者之间的

〔60〕 See Christopher Kuner et al., *The EU General Data Protection Regulation (GDPR) A Commentary*, Oxford University Press, 2020, p. 523, 666.

〔61〕 See A29WP: A29 WP, Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679, 17/EN. WP 251rev. 01 (Feb. 6, 2018).

〔62〕 参见前引〔46〕，最高人民法院民法典贯彻实施工作领导小组主编书，第253页。

〔63〕 参见杨显滨：《网络平台个人信息处理格式条款的效力认定》，载《政治与法律》2021年第4期。

关联性。^{〔64〕}可能面临的困境在于，打开“算法黑箱”不可避免地涉及对定义代码和算法规则的专业解释，高维机器学习人类解释风格在技术素养方面的差异造成了算法透明度的一大障碍。^{〔65〕}此时需要由该自动化决策所涉及的具体领域的专业人士作出阐释，如医疗服务者可以解释手术与辐射治疗的差异，对每个技术细节进行解释，^{〔66〕}以辅助判断决策作出的正当性。

（二）损害赔偿认定中的算法解释

事后履行解释义务所发挥的效果弱于事前解释的防御效果，需要辅之以拒绝、删除，以及合同和侵权领域的损害赔偿责任。《个人信息保护法》第24条第3款赋予了个人信息主体对仅通过自动化决策作出决定予以拒绝的请求权，主张拒绝的要件应为该决定对个人权益产生了重大影响。因此，个人信息主体有权要求个人信息处理者予以解释说明，在个人信息处理者违反解释义务时，个人信息主体可主张拒绝仅受该自动化决策结果约束，亦可根据《个人信息保护法》第47条第1款第4项的规定主张删除。但“拒绝+删除”的处理模式仅能面向将来产生保护效力，并不能有效救济已经遭受重大影响的个人权益。

存在重大影响但未造成实际损害的情形下，个人信息主体仅可主张算法解释，存在重大影响且损害已经实际发生的情形则需在主张算法解释后，在个人信息处理者未对该行为作出合理解释的情况下才能主张损害赔偿。是否能够辅之以合同或者侵权责任意义上的损害赔偿请求权，关键在于判断重大影响的结果是否能够被评价为法律意义上的损害。^{〔67〕}依《个人信息保护法》第24条第1款的规定，个人信息处理者不得对个人在交易价格等交易条件上实行不合理的差别待遇。此种差别待遇构成价格歧视，在合同关系领域可能构成欺诈或者显失公平，其损害的是个人信息主体的财产权益。但该条第2款所提及的信息推送、商业营销等针对个人特征的选项若未采取滥用市场支配地位、故意悖俗地使用算法造成排除、限制竞争或者损害消费者福利等行为时，不可被直接认定为对个人信息主体构成了损害。因此，是否能够主张合同或者侵权责任上的损害赔偿请求权，需根据《民法典》第584条或第1182条至1185条的规定具体判断。

“未合理履行算法解释义务”这一行为在合同法上与侵权法上的评价存在区别。在传统民法中，侵权行为与违约行为均为债的发生原因，前者产生法定之债，后者产生意定之债。^{〔68〕}从用户角度出发，数据活动关系人可通过用户协议（个人信息授权合同）的方式建立数据收集、利用的债之关系。^{〔69〕}如在定制信息推送、智能投顾、搜索引擎等典型场合，算法提供者的主给付义务在于为个人提供上述核心服务功能，而算法解释旨在避免算法决策对当事人的个人权益产生重大影响，可被合同关系中的保护义务所涵盖。^{〔70〕}算法解释义务在合同领域表现为对合同履行行

〔64〕 See GDPR Article 15 (1) (h).

〔65〕 See Bryce Goodman, Seth Flaxman, European Union Regulations on Algorithmic Decision-Making and a Right to Explanation, 38 *AI Magazine* 50, 55 (2017).

〔66〕 See Tae Wan Kim, Bryan R. Routledge, Informational Privacy, A Right to Explanation, and Interpretable AI, in 2018 *IEEE Symposium on Privacy-Aware Computing (PAC)*, IEEE, 2018, pp. 64–74.

〔67〕 参见王泽鉴：《损害赔偿》，北京大学出版社2017年版，第62页。

〔68〕 参见程啸：《侵权责任法》（第3版），法律出版社2021年版，第79页。

〔69〕 参见龙卫球：《数据新型财产权构建及其体系研究》，载《政法论坛》2017年第4期。

〔70〕 对于“给付义务—保护义务”这一分类，参见张家勇：《合同法与侵权法中间领域调整模式研究——以制度互动的实证分析为中心》，北京大学出版社2016年版，第199–200页。

为符合双方缔约目的的解释义务。若个人信息处理者未能解释对个人权益产生重大影响的自动化决策条款或者自动化决策结果,则分别导致该条款未被订入合同或个人信息处理者承担相应违约责任的后果。有观点认为算法服务提供者的算法解释属于合同法上的从给付义务与附随义务,^[71]实际上也是承认算法解释作为合同义务保护当事人合法权益的本质。

因此,当个人信息处理者与个人信息主体处于合同关系中时,“未合理履行算法解释义务”应被评价为债务不履行。当事人将“自动化决策决定对个人权益产生重大影响则可要求个人信息处理者作出解释”订入合同条款并不能将解释义务的性质划定为合同的主给付义务。原因在于,合同的主给付义务内容为个人信息主体按要求提供相关个人信息并获得服务,个人信息处理者依据个人信息主体提供的个人信息作出自动化决策并提供服务。但算法解释义务与主给付义务内容密切相关,其性质为合同保护义务,对该义务的违反亦构成主给付义务的不适当履行,个人信息主体可据此主张违约损害赔偿。

在个人信息处理者与个人信息主体并未进入合同这一特别结合关系时,“未合理履行算法解释义务”不可被评价为侵权法上的“侵权行为”,而应被评价为“过错”。对于侵权请求权中“过错”要件的举证需通过算法解释辅助说明,如个人信息处理者无法解释或者拒绝解释则应推定个人信息处理者存在过错。侵权行为所违反的是法定义务、绝对义务,^[72]侵害的对象一般为绝对权。算法解释义务作为人格权所衍生的非独立性请求权,并不具有绝对义务的性质。违反算法解释义务本身并非侵权行为,对个人权益造成重大不利影响的个人信息处理才属于个人信息保护中的侵权行为。需要注意的是,《个人信息保护法》第24条第3款所规定的“个人权益”并不仅限于个人信息权益,亦包括其他权益。典型如算法处理个人信息过程中对个体的信用评级造成重大不利影响,侵害了信用权(《民法典》第1029条)。因此,应将“未合理履行算法解释义务”评价为过错,但需要厘清的是归责原则为一般的过错责任原则还是过错推定责任,这与举证责任的分配以及不利后果的承担密切相关。

本文从民法体系出发,将个人信息处理关系限定在采用自动化决策的非公务机关与个人信息主体之间。在此种关系类型中,双方当事人在举证责任和诉讼能力上的差异主要是由“自动化决策技术”造成的,若采用一般过错归责原则,鉴于自动化决策技术中蕴藏的算法复杂性,可能会对个人信息主体课加过重的举证负担,^[73]因此应采过错推定原则。^[74]《个人信息保护法》第69条第1款规定了个人信息权益致害以过错推定原则归责,^[75]处理个人信息侵害个人信息权益造成损害,个人信息处理者不能证明自己没有过错的,应当承担损害赔偿等侵权责任,目的在于维持权益保护与产业发展的平衡。^[76]虽然个人信息侵权中的损害具有无形性、潜伏性、未知性等

[71] 参见鲁春雅:《自动化决策的算法解释权——以合同义务为视角》,载《中国社会科学报》2020年12月23日,第4版。

[72] 绝对义务是指个人具有与不定数目的其他人或其他所有人有关的义务。参见〔奥〕凯尔森:《法与国家的一般理论》,沈宗灵译,商务印书馆2013年版,第140页。

[73] 有法院认为:“从收集证据的资金、技术等成本上看,作为普通人的庞理鹏根本不具备对东航、趣拿公司内部数据信息管理是否存在漏洞等情况进行举证证明的能力。”北京市第一中级人民法院(2017)京01民终509号民事判决书。

[74] 参见叶名怡:《个人信息的侵权法保护》,载《法学研究》2018年第4期。

[75] 参见前引〔32〕,程啸书,第506页。

[76] 参见谢鸿飞:《个人信息泄露侵权责任构成中的“损害”——兼论风险社会中损害的观念化》,载《国家检察官学院学报》2021年第5期。

诸多与传统侵权场域不同的特殊性，^{〔77〕}但在自动化决策处理个人信息的场合，关键在于证明算法规则具有合理性且未对个人权益产生重大不利影响，以算法解释作为评价个人信息处理者过错的衡量标准。“不能证明自己没有过错的”不仅包括个人信息处理者履行解释义务后无法证明决策结果合理性的情形，亦包括个人信息处理者拒绝履行算法解释义务的情形，拒绝解释视为放弃举证而直接负担不利后果。

值得讨论的是，是否需对敏感个人信息与非敏感个人信息的归责原则作区分处理。有观点认为处理敏感个人信息应适用无过错责任，处理非敏感个人信息适用过错推定责任，原因在于敏感个人信息的处理行为开启了危险源，处理者对处理行为具有控制力，应承担更严格的责任。^{〔78〕}在本文的语境下，个人信息处理为私主体之间的关系，并不包括公务机关，双方对于危险控制的差异在于算法技术的控制权，采用过错推定责任已经对个人信息处理者设定了自律机制，以增强其处理过程中的责任心。^{〔79〕}对敏感个人信息与非敏感个人信息的区别保护可通过调整算法解释义务即举证责任的方式实现。在敏感个人信息处理的场合，个人信息处理者的举证责任更重，原因在于敏感个人信息相较于一般个人信息更容易在处理过程中受到侵害，对算法规则设计的合理性要求更高。若在敏感个人信息处理中适用无过错原则，相当于将处理敏感个人信息的行为直接视为危险责任，既是对算法解释义务作为过错评价功能的忽视，亦不利于解决个人信息流通与个人信息保护之间的利益冲突。综上，若个人信息处理者拒绝履行解释义务或者履行解释义务后并不能证明其无过错，则损害赔偿的不利后果由其承担。至于损害赔偿的范围，依《个人信息保护法》第69条第2款具体确定，赔偿范围包括财产损失与精神损害。

四、结 论

现代性正以前所未有的方式，把我们抛离了所有类型的社会秩序的轨道，从而形成了其生活形态。这一点在技术方面表现得尤为显著，并逐渐渗透至其他领域，引发信任与风险的紧张关系。^{〔80〕}个人信息应作为人格中纯粹的一部分，不可简单地沦为促进他人增益或者实现社会目的的手段。从民法体系的视角观察，算法解释背后所保护的个人信息权益应为人格权益，算法解释请求权的本质是为人格权服务的非独立请求权。在合同关系领域，算法解释可纳入合同保护义务的范畴以助于主给付义务的履行，个人信息处理者可能因违反解释义务承担相应的违约责任。在侵权领域，算法解释则作为过错评价的要素发挥作用，即自动化决策对个人权益产生重大影响可否被评价为侵权行为，需以个人信息处理者履行算法解释义务的结果作为辅助判断的标准。

为缓解个人信息权益保护与个人信息利用之间的紧张关系，算法解释应区分为事前解释与事后解释两种类型。事前解释仅适用于敏感个人信息的自动化决策场合，因为在事前无法对一般个人信息是否产生重大影响作出准确的判断，而敏感个人信息的处理基于其个人信息类型的敏感度

〔77〕 参见田野：《风险作为损害：大数据时代侵权“损害”概念的革新》，载《政治与法律》2021年第10期。

〔78〕 参见程啸：《侵害个人信息权益的侵权责任》，载《中国法律评论》2021年第5期。

〔79〕 参见申卫星：《论个人信息保护与利用的平衡》，载《中国法律评论》2021年第5期。

〔80〕 参见〔英〕安东尼·吉登斯：《现代性的后果》，田禾译，译林出版社2011年版，第4-8页。

即可推定出可能会对个人权益产生重大影响。重大影响实际产生后的事后解释则无需区分个人信息的敏感程度。此时若产生了损害,单纯的事后解释与“拒绝+删除”模式并不能发挥全面的救济功能,需辅助以合同法或者侵权责任上的损害赔偿赔偿责任。是否承担合同损害赔偿赔偿责任取决于是否违反合同义务。在侵权损害赔偿中,若个人信息处理者无法解释或者拒绝解释则推定其存在过错并承担不利后果。

Abstract: From a civil law perspective, personal information should be protected as a legal interest. The algorithmic interpretation of automated decision-making is in fact a non-independent claim derived from personality rights and does not require the creation of a separate right of algorithmic interpretation. In the field of contract, algorithmic interpretation can be included in the scope of the duty of protection, while in the field of tort, it has the function of fault evaluation, and when the personal information processor is unable to explain or refuses to explain, it can be presumed to be at fault. In order to protect the commercial secrets of personal information processors and reduce their burden of interpretation, the algorithmic interpretation stipulated in Article 24 (3) of the Personal Information Protection Law shall be distinguished into two types: ex ante interpretation and ex post interpretation, which differ in terms of the construction of the elements and the degree of interpretation. The ex ante interpretation shall be limited to the handling of sensitive personal information that “may have a significant impact”, while the ex post interpretation shall occur in the case of automated decisions that “already have a significant impact on the rights and interests of individuals”. The duty of ex ante interpretation should be lower than that of the ex post interpretation, which corresponds to an abstract interpretation of the system’s function and a detailed interpretation of a specific decision respectively. In the event of actual damage arising from the material impact, the subject of personal data may claim liability for damages in breach of contract or tort. Intentional exemptions from the obligation to interpret ex ante shall be void if they violate Article 497 or Article 506 of the Civil Code, but the obligation to interpret ex post and the right to claim damages arising therefrom may be exempted.

Key Words: algorithmic interpretation, interest in personal information, automated decision-making, significant impact

(责任编辑:殷秋实 赵建蕊)