数据推理的法律规制

苏宇[1]

【摘 要】：数据推理既是开发利用数据价值的重要手段，也可能对个人信息保护、数据安全和算法安全造成潜在风险。规制数据推理风险既需要尽量追求数据安全与数据开发利用之间的价值平衡，也需要有充分的技术前瞻性。国内相关立法中尽管已有较多间接支持和保障数据推理规制的规范依据，数据推理规制的直接依据及具体制度工具仍颇为有限。在众多专门性法律机制中，应当以元规制为主要的数据推理规制路径，建立和完善以数据分析安全合规为重心的规制体系。

【关键词】：数据推理 数据安全 个人信息 元规制

数据推理既是开发利用数据价值的重要手段，也可能对个人信息保护、数据安全和算法安全造成潜在风险。伴随《数据安全法》和《个人信息保护法》等法律规范的起草、制定和实施，数据推理的法律问题开始受到重视。数据流通与利用主要方式的变迁，更使数据推理在数据治理中的关注度日益提升。随着数据治理立法的持续推进与相关执法活动的不断加强，直接窃取具有一定敏感程度的原始数据(包括个人信息与非个人信息)将日益困难，此时基于“非接触性”的方式间接推测上述数据就成为主要的风险来源，这就使数据治理必须充分注意来自数据推理的威胁。

在我国，以个人信息相关法益为代表的权益保护要求使数据的流通与交易越来越倾向于限制原始数据的流通与转移，而代之以提供被处理数据的数据利用方式。2022 年 1 月，国务院办公厅印发的《要素市场化配置综合改革试点总体方案》在“建立健全数据流通交易规则”部分明确提出“探索‘原始数据不出域、数据可用不可见’的交易范式”。在相关法律和政策的推动下，隐私计算相关技术在业界可谓“热浪滔天”,隐私集合求交、样本对齐、零知识证明、同态加密、差分隐私、安全多方计算、可信执行环境乃至全域隐私计算等术语及相关技术的出现和流行，使得海量数据可以在“可用而不可见(知)”的前提下经由一定的数据分析过程而被开发利用。

以北京国际大数据交易所为代表的“数据可用不可见，用途可控可计量”新型交易范式[1]很有可能在未来一段时间主导数据交易及后续开发利用的进行。原始数据“不可见”的要求显然将很大程度阻断外界非法接触原始数据的途径。在基础性管理制度日臻完善、直接窃取原始数据越来越困难的前提下，数据安全风险将逐渐转入隐蔽的疆域。原始数据的安全和个人信息保护的成效，最终取决于数据推理所能到达的边界。通过各种途径获取数据的主体能够通过数据推理在多大程度上推断原始数据或推知其中隐含的敏感信息，不仅很大程度上决定着《个人信息保护法》中“去标识化”与“匿名化”之间的分野，更在很大程度上决定了对数据安全保护技术的需求强度，进而间接决定了数据安全与数据价值利用之间的潜在平衡点，可谓未来数据开发利用的要害所在。

由此，数据推理的法律规制也需要学界与实务界更多的关注。数据推理在法律层面上的重要性不仅在于其对数据安全和个人信息保护日益举足轻重，更在于其技术上的可能边界与业务上的可行范围很大程度上影响着数据治理中相关权益的配置和制度设计。数据推理的可能性将直接影响数据分类分级的结果，而法律如何防止潜在的攻击者利用数据推理获取本来无从知悉的

[1]**作者简介**：苏宇,中国人民公安大学法学院副教授、博士生导师,中国人民公安大学数据法学研究院院长。主要研究领域为行政法学、数据法学。中国人民大学法学学士,北京大学法学硕士、法学博士,美国加州大学伯克利分校访问学者,中国法学会网络与信息法学研究会理事。曾在《中国法学》《法学研究》等学术刊物发表论文 50 篇,9 篇次被《新华文摘》《中国社会科学文摘》、人大报刊复印资料转载或摘编。主持或参与国家级、省部级课题 20 余项。曾获第六届全国教育科学研究优秀成果奖三等奖、第八届“董必武青年法学成果奖”二等奖。
基金项目：国家社会科学基金重大项目“网络信息安全监管法治研究”(21&ZD193)研究成果

敏感信息，则很大程度上影响着数据治理的成效，法律规制数据推理的方式和深度，更将深刻影响数据要素的价值实现和数字经济的发展前景。

一、数据推理的多重风险与规制需求

在宽泛意义上，数据推理旨在描述数据之间明确或不明确的数据关联或数据联系。[2]利用数据进行推理的具体方式可谓五花八门，例如基于粒计算(包括词计算、粗糙集、三支决策等)的推理，利用机器学习算法进行知识发现，基于结构方程模型的因果推断，基于信息熵函数的贝叶斯因果推理，等等，尚未形成统一的数学基础及技术路线。无论数据推理的具体技术路线及实施过程如何，总体上，各种推理均致力于通过利用已有数据中包含的信息，分析数据之间的相关关系或/和因果关系，进而推导或推测出未知的信息。在法律上，数据推理从属于《数据安全法》第 3 条界定的“数据处理”,后者的含义极为宽广，“包括数据的收集、存储、使用、加工、传输、提供、公开等”的表述几乎涵盖了所有与数据有关的活动。其中，数据推理是使用数据的重要方式，其对于数据开发利用的价值日益显著：从数据中能够推测出多少有用的信息，决定了数据潜在的开发利用价值。

然而，数据推理的风险也不容忽视。信息安全领域的研究者早已注意到一个基础性的风险：“不同的信息之间有着内在的联系，这些联系可以使得用户可以根据合法获得的信息推理出不可访问的敏感信息。”[3]基于此种可能性，数据推理可能对个人信息、数据安全直至算法安全等造成威胁。尽管数据推理引致的多重风险未必能与其为数字经济和数据资源利用所带来的巨大利益相提并论，必要的法律规制也可以使数据推理更好地服务于正当的数据利用目的，但准确认识数据推理的规制需求非常关键。

(一)数据推理的多重风险

数据推理最容易为公众所感知的风险是对个人信息法益的侵害。随着个人信息保护需求逐渐凸显，不少学者已经认识到基于聚合数据的推理暴露个人信息的可能性。早在 2011 年，就已有学者明确提出数据聚合与分析能拼合有关个体的数据，形成更丰富的个体画像。[4]在个人信息界定标准的相关研究中，数据推理的相关内容已经不时可见：一项信息能否与其他信息结合完成指向特定自然人的推理，意味着其是否满足“间接识别”的标准；[5]一项去标识化的信息是否能经由推理还原信息主体的相关属性，已成为判断去标识化与匿名化的分水岭，即判断是否符合匿名化标准的重要考量因素。[6]数据推理的可能性边界很大程度上决定了间接识别或再识别的概率，从而使得这一技术问题开始具备日益显著的法律意义。

在经由数据推理实现“再识别”的潜在压力下，单纯的“去标识化”已不足以实现数据资源开发利用中个人信息保护的要求，许多场景下必须实现匿名化处理。由此，一系列技术标准正在陆续形成，单有推荐性国标《信息安全技术个人信息去标识化指南》(GB/T37964-2019)及团体标准《个人信息处理法律合规性评估指引》(T/CLAST001.2—2021)等还不足以支撑完整的个人信息保护需求，全国信息安全标准化技术委员会正在制定的《信息安全技术个人信息去标识化效果分级评估规范》将使个人信息保护与数据开发利用走向更为精细的平衡，其中决定性的因素就是重标识风险是否低于阈值，即去标识化后的个人信息能否经受常规的数据推理之冲击。

即便如此，个人信息仍然难以免遭数据推理之威胁。例如，推理者可以利用“脸书”(Facebook)的点赞情况等行为踪迹数据自动化地推测信息主体的年龄、性别、宗教信仰、受教育程度等个人信息。[7]由此，不仅未匿名的信息主体容易暴露更多个人信息，已被认为是经匿名化处理的个人行为踪迹数据集如被推测出足够丰富的画像信息，也有可能在特定情形下实现再识别。

不仅个人信息可能蒙受数据推理的侵害，数据安全更有可能直面数据推理的威胁。拼合“用户画像”之类的推理甚至并非数据推理的主要风险，数据推理最集中的威胁目标是具有高度敏感性的商业信息乃至政治、军事信息，此种威胁随着大型数据库的诞生而愈发明显。长期以来，数据库推理控制一直是数据库安全领域的重要研究内容，数据库推理控制的研究早在 20 世纪 70 年代末就已开始，这是因为对于数据库的数据安全而言，推理攻击的风险是始终存在的。“在多级安全数据库中，若低安全级别的用户利用其已知的数据关联性，及其有权直接读取的低安全级别数据，推理出更高安全级别的数据信息，就构成对数据库的推

理攻击，造成高安全级别数据的泄漏。”[8]

20 世纪 90 年代是数据推理研究的“黄金时代”,研究者发现，数据库中的知识发现可以通过引入传递外部知识的“催化关系”来增强，此种“催化推理”能催化新的推理通道，虽然其本质上是不精确的，但推理的粒度可能足够精细，也足以造成安全隐患，并且还可以利用搜索引擎由机器进行自动化的推理。[9]无论是基于敏感数据的查询推理(将不同查询的答案加以组合),还是基于数据与元数据(Metadata)的结合，都有可能推断出敏感信息。[10]在知识推理伴随人工智能技术的发展而广泛兴起后，数据推理对数据安全的威胁显著升级。基于描述逻辑的推理、基于图结构和统计规则挖掘的推理、基于知识图谱表示学习的推理、基于神经网络的推理和混合推理等众多推理方式使得知识推理有机会发现更深层的数据关联性，[11]这就使数据安全陷入更深层的风险之中。即便采取了隐私计算技术也未必能保证免受数据推理攻击之侵害，例如数据推理攻击可以从联邦学习的模型参数或梯度中反推原始数据，从而对隐私计算技术所保护的数据安全造成威胁。[12]

数据推理对于算法安全的影响在人工智能时代也逐渐体现出来。“深度学习算法本质上是在所有概率测度构成的空间内进行优化，算法既记住了训练样本，又学习了内在知识。”[13]既然算法模型中已经融合了训练样本的信息，攻击者就可以通过复杂的推理方法(如在训练影子模型的基础上进行推理)攻击人工智能系统以窃取其训练数据，即发动所谓“成员推理攻击”(membership inference attacks),对用户隐私和数据安全造成显著威胁。[14]最著名的实例是大型人工智能算法模型 GPT-2 因为过拟合(overfitting)而被攻击者反推其中用于训练的个人数据。[15]

如果仅仅是窃取训练样本，算法安全面临的风险与数据安全仍可谓基本重叠，问题是通过训练影子数据集和影子模型之类的方式还可以基于计算设备在运行深度神经网络时产生的侧信道信息(如功耗)等进行量化推理，进而提取模型结构和权重等算法模型中的关键信息，威胁算法安全。[16]另外，利用训练“冒牌模型”(knock-offmodel)等类似的方式，还可以在不获取任何模型参数与结构特性的前提下，从输入与输出数据的关系中推理出算法黑箱的某些特性，从而“窃取”其性能并威胁算法模型的安全。[17]

质言之，数据推理可以基于数据之间的各种相关关系或因果关系，发现或估测数据中难以被直接发现或直观认知的隐含信息。受威胁的目标信息不仅仅是敏感的原始数据本身，也有可能是相关数据的某些关键数学特征，如数据流形维度、概率分布、稀疏度等；推理的对象还包括上述特征的动态变化，因为敏感数据的数学特征如果发生某种结构性的变化，即便原始数据不泄露，其背后所隐含的秘密(如军事行动)也可能被发现。这些隐含信息一旦被推理者推知，其隐含的法益也将直接或间接面临被侵害的风险，不同场景下对数据推理的规制需求亦由此滋生。

(二)数据推理的规制需求

尽管数据推理的规制需求日益凸显，学界与实务界仍然对此保持非常谨慎的态度。这是因为对数据推理的规制一旦失当，很有可能波及对数据价值的正常开发利用。在《个人信息保护法》和《数据安全法》制定以前，学界对数据推理的规制需求罕有关注。随着数据开发利用进程的不断发展，来自数据推理的威胁正在日益凸显数据安全目标与网络安全目标的差异：以网络安全等级保护为代表的网络安全制度不能涵盖数据安全和个人信息保护的规制需求，因为推理者理论上完全可以基于依法获取的数据开展推理，在不涉及违法网络活动的前提下威胁到数据安全和个人信息保护的目标。

这一问题背后隐含着数据安全的双重向度：网络安全主要考虑的是“为”的向度，即网络行为是否正常、能否避免或阻止对网络的各种破坏或入侵行为；数据安全不仅要考虑“为”的向度(行为向度),还要考虑“知”的向度(信息向度),即数据中所隐含的信息是否会被应知范围以外的主体获悉。这意味着即便没有发生任何网络攻击或其他形式的恶意入侵、没有发生任何天灾人祸和设备故障，数据安全也可能受到威胁。[18]

正因如此，数据推理所产生的负面后果可能非常隐蔽：无论是对个人信息和敏感数据的提取，还是对算法模型的破解，都可

以在外界难以觉察的情况下完成。通过设置网络安全保护义务等传统的法律规制方式已难涵盖数据推理所导致的法益侵害风险，治理者需要运用新的规制手段，在法治框架内实现数据推理风险的精准治理，有效平衡数据开发利用与防范数据推理风险的目标。对此，两个关键的规制需求是必须注意的：

第一，规制数据推理风险需要尽量追求数据治理的价值平衡，避免导致数据开发利用方式的进一步收缩。数据安全保护程度的提高与数据价值的充分开发利用之间在一定范围内存在此消彼长的关系，尽管难以量化计算最优的精准平衡点，也需要在整体上强调规制目标上的平衡取向。数据的流通交易方式从单纯的去标识化(例如采用 xID 标记技术替换原 ID 但保证映射关系)到“原始数据不出域、数据可用不可见”的转变，已经使数据的利用价值一定程度上让步于个人信息保护和数据安全等目标；而在“数据可用不可见”的技术谱系中，还存在安全性与数据利用价值的强弱转换关系，不能因为一味追求完全不发生数据推理的风险而过度削减流通数据所包含的信息量。例如，全同态加密后的数据在抵抗推理方面有相当强的安全性，但被全同态加密的数据能实现的利用价值尚较为有限，而且全同态加密的成本和效率足以让普通数据利用者望而却步。[19]又如，抑制技术(suppression techniques)采用删除部分数据项的方式避免重标识化风险，噪声添加技术可以使数据在改变原始值的同时保持部分统计特性，泛化技术可以降低数据集中所选属性的粒度，这些技术一定程度上都可以防御数据推理的威胁，但同时也不可避免地(甚至是大幅度地)损害数据的效用。质言之，“一刀切”式的粗放监管“无助于解决数据流动与数据安全之间的内在冲突”,20 数据推理的规制应当尽可能通过避免损害数据效用的方式进行，追求以较低的成本和负担实现规制目标。

第二，规制数据推理风险需要有充分的技术前瞻性。数据推理的能力边界取决于数学理论与数据技术的发展。数学理论与数据技术的发展呈现不平衡、不确定、信息不对称的特性，对数据推理的规制构成了实质性的挑战。“不平衡”是指不同国家、企业、高校、研究机构及学者在相关数学原理及数据技术的发展上相去甚远，前沿技术的重大发展高度集中于产业巨头、著名高校及头部科研机构，数学原理上的突破则甚至可能仅为极个别学者所掌握和理解。“不确定”是指数学原理和数据技术的重大发展无法被精确预测，何人于何时何地以何种方式及程度完成突破性研究进展，即便是同行专家也很难准确估计。“信息不对称”是指由于商业秘密保护、交流障碍及知识基础差距等原因，一般研究者及从业者在数学原理与数据技术方面往往难以及时充分了解上述重大突破的完整内容，即便能接触、理解和消化，也会存在明显的时滞和信息差。“信息不对称”也可以被理解为一种颇具威胁的隐蔽性：即便完全满足数据分类分级保护的标准，基于深入的相关性分析及因果推断，在占有足够体量及维度的低敏感度数据基础上，专业分析者依然有可能悄然获得背后潜藏的特殊敏感信息。21 数据推理规制必须充分考虑上述特点的影响。如缺乏合适的规制工具和制度设计，规制者将很可能长期处于被动应对数据推理技术最新发展的处境，需要规制者在建设治理机制时保持相当程度的前瞻性。

这两项特殊的规制需求使得对数据推理的规制颇为棘手，当前国内外相关法律规范在此方面也仅有相对原则性的一般要求和较为分散的专门性法律机制，尚未形成数据推理规制的系统性法律方案。此种局面固然不利于数据推理规制的全面实施，但也为探索合理、灵活、专业的规制进路与方案留出了必要的空间。

二、数据推理规制的规范依据

目前，国内外有关数据推理规制的依据散见于数据治理、个人信息保护及消费者权益保护等领域多个层次的规范中。相关规定尽管在具体调整对象和效力位阶上不尽一致，但主要内容较为接近，体现了立法者对数据推理规制在基本取向上的价值共识。以国内立法中数据安全和个人信息保护方面的相关规定为例，数据推理规制的规范依据已初具雏形，开始形成支持数据推理规制的制度合力，但其往往尚不具备充分的可操作性，更多地起到价值导向的作用。

(一)数据安全法制中的相关规定

《数据安全法》中没有直接规制数据推理的规定，但部分条款为相关配套法规规章的制定及数据合规体系的建设奠定了基础。《数据安全法》第 8 条对数据推理活动规定了总体上的合法性边界。《数据安全法》第四章“数据安全保护义务”针对所有开

展数据处理活动的主体及若干特殊主体规定了一系列义务。上述规范尽管并不单独针对数据推理活动，但至少为数据推理活动奠定了建立合规体系的基础。此外，《数据安全法》第 17 条有关数据开发利用技术和数据安全标准体系建设的规定、第 18 条有关数据安全检测评估及认证的规定以及第三章“数据安全制度”中数据安全应急处置机制、数据安全审查制度等多项规定，已经为包括数据推理规制在内的数据安全治理建立了基础性的框架，为下位法中相关的专门性规定乃至专门性的配套立法提供了有力支持。

在《数据安全法》的基础上，各地方、各部门有关数据治理的立法中已陆续出现若干与数据推理有关的专门性规定。除一般性地要求避免泄露敏感数据或加强风险管理的条款外，以下几方面的专门性规定值得关注：

一是数据利用目的与用途的限制。例如，2019 年制定的《贵州省大数据安全保障条例》第 9 条明确了“意图合规”的原则，第 22 条又规定：“使用数据不得用于非法目的和用途。明知是通过攻击、窃取、恶意访问等非法方式获取的数据，不得使用。”2020 年制定的《浙江省公共数据开放与安全管理暂行办法》第 22 条第 1 款则确定了约定利用范围的用途限制。

二是(大)数据分析范围的外延式界定。此方面的法律规范暂时仍未见实例，但 2021 年国务院办公厅印发的《要素市场化配置综合改革试点总体方案》(国办发[2021]51 号)已提出“探索建立数据安全使用承诺制度，探索制定大数据分析和交易禁止清单，强化事中事后监管”。

三是数据处理全流程记录、身份认证、访问控制、安全审计、安全保护协议、数据攻防演练等有助于间接限制数据推理的数据安全机制，例如 2021 年制定的《深圳经济特区数据条例》在第 75 条规定：“数据处理者应当对其数据处理全流程进行记录，保障数据来源合法以及处理全流程清晰、可追溯。”《贵阳市大数据安全管理条例》第 19 条规定：“市人民政府建立大数据安全靶场和产品检验场地，对大数据安全新技术、新应用、新产品进行测试、检验，定期开展攻防演练，促进大数据安全城市建设。”这些制度并不直接限制数据推理，但增强了发现、预防乃至阻止滥用数据推理的能力。

不仅如此，一些必要的控制措施(如权限控制)也开始在规范性文件中出现。例如 2020 年制定的《广西公共数据开放管理办法》(桂数发[2020]23 号，属规范性文件)第 27 条第 2 款规定：“各地各部门各单位如发现数据使用主体对获取的公共数据出现未授权使用、非法滥用、未经许可的扩散和泄露等行为时，应及时关闭数据使用权限。必要时，应依据有关法律法规追究其法律责任。”总体上，数据安全法制已经为数据推理规制的进一步展开建立了初步的制度基础。

(二)个人信息保护法制中的相关规定

与数据安全法制类似，《个人信息保护法》中亦无直接针对性的条款，但也对相关的数据推理活动规定了一般性的目的限制与合法性边界。例如，《个人信息保护法》第 6 条第 1 款、第 7 条对于规制滥用个人信息进行推理的行为有着基础性的价值指引意义。此外，《个人信息保护法》第 17 条的告知、第 19 条的最短保存期限、第 22 条和第 23 条的重新取得同意、第 27 条的处理公开个人信息之限制、第 28-32 条对敏感个人信息的处理规则等等，都可以间接限制基于个人信息的数据推理活动，降低数据推理侵害个人信息相关法益的风险。

尽管《个人信息保护法》中缺乏直接规制数据推理的条款，不少地方立法及部门行政立法对此已有所作为。例如，2021 年制定的《深圳经济特区数据条例》第 11 条第(五)项规定：“……建立最小授权的访问控制策略，使被授权访问个人数据的人员仅能访问完成职责所需的最少个人数据，且仅具备完成职责所需的最少数据处理权限。”这一规定符合数据安全的基本要求，可以很大程度上限制基于个人信息的数据推理之滥用。[22] 更直接的规定亦不乏实例，如 2017 年制定的《江苏省预防未成年人犯罪条例》第 35 条第 3 款规定：“涉及欺凌和暴力事件的报道、信息，不得泄漏未成年学生的姓名、住所、照片以及可能推断出未成年学生信息的资料。”类似规定在预防未成年人犯罪、统计管理、卫生健康、人口调查等方面的地方立法中已不鲜见。2020 年制定的《深圳经济特区健康条例》第 85 条及第 87 条第 1 款甚至致力于从正反两方面平衡数据开发利用和数据推理风险预防。

不过，这些规定基本上是零星出现在不同层级和领域的规定中，且绝大多数仅有含义宽泛的一般禁止性要求。

数据安全和个人信息保护法制中与数据推理有关的相关规定清晰地凸显出亟待完善的制度建设方向：尽管可以间接支持和保障数据推理规制的规范依据已较为丰富，但这些规定尚不具备充分的可操作性，难以具体指引数据推理规制的展开；数据推理规制的直接依据及具体的制度工具仍颇为有限，颗粒度粗、针对性低且重心不明的规制框架恐难精细地掌握数据开发利用与数据推理风险预防日益复杂的价值平衡。因此，对数据推理的规制需要建立更加直接、精准的专门性法律机制，其中部分机制已有一定的研究或实践基础。

三、数据推理规制的专门性法律机制

数据推理规制在技术层面已有一定积累。经过多年业务实践，许多数据库的开发者早已意识到需要对数据库进行推理控制以确保数据分析安全，推理通道控制已成为数据库安全研究和实践的常见内容。这一概念的首创者斯达顿(Jessica Staddon)指出，在多级数据库中，如果用户能够从一系列低安全等级的查询中推断出敏感信息，则这些查询共同构成了一条推理通道(Inference Channel)。[23]随着以推理通道控制为代表的数据库推理安全技术不断发展完善，在一个多级数据库内部控制数据推理滥用的管理机制和安全措施逐渐成型，并开始进入相关技术标准。

在我国,《信息安全技术数据安全能力成熟度模型》(GB/T37988-2019,业界常简称“DSMM”)中首次对“数据分析安全”提出了专门性、系统性的具体要求，为数据安全工作提供了重要的视野和认知维度。不过，对基于结合数据库内外的数据而进行的推理活动，技术上难有风险预防的万全之策。尽管如此，对于保障正常利用数据的前提下尽量降低滥用数据推理风险的规制需求，来自技术层面的局部支持仍是必要且可行的。基于现有技术路线，数据推理规制的专门性法律机制可以区分单一数据源及复杂数据源的情形进行讨论。

(一)单一数据源的数据推理规制

由于数据库安全领域有关推理控制的标准渐趋成熟，单个数据库(包括单一数据库和多重数据库)的推理控制已经几乎不需要技术标准以外的法律规制工具。近 30 年来，数据推理安全研究已经从统计数据库推理控制扩展到数据仓库推理控制，细化了推理控制的粒度，强化了推理控制的动态性。[24]在世界范围内，从 20 世纪 80 年代开始，相应技术标准亦随之得以陆续形成和完善，为单一数据库甚至数据仓库的推理规制提供了有益助力。只要所有操作都在数据库内部进行，相应的技术性监管工具乃至技术标准即有运用之空间。

在国内，早在 2002 年,《计算机信息系统安全等级保护数据库管理系统技术要求》(GA/T389-2002)就已专节(A.8)规定推理控制的要求，除归纳了数据推理的方法和用于推理的信息类型外，还简要地归纳了多级数据库的基本安全原则和防止推理的方法。[25]这一要求亦见诸其后制定的推荐性国标《信息安全技术数据库管理系统安全技术要求》(GB/T20273-2006),但这一标准后来被新的同名标准《信息安全技术数据库管理系统安全技术要求》(GB/T20273-2019)完全取代。在新标准中，有关推理控制的部分被删除，同时规定“对该部分内容的评估直接依据 GB/T18336.3—2015 各级别的脆弱性评定要求进行评估”。在其提及的《信息技术安全技术信息技术安全评估准则第 3 部分：安全保障组件》(GB/T18336.3—2015)中，针对数据推理的脆弱性评定要求是由“脆弱性分析”保障组 AVA_VAN 处理的，不同安全等级的数据库需要不同程度考虑隐蔽信道，并能够经受相应程度的穿透性测试。在同一归口单位负责并由基本相同的主体起草、专门执行评估的标准《信息安全技术数据库管理系统安全评估准则》(GB/T20009-2019)中，对穿透性测试和子集信息流控制有进一步展开的详细要求。

在数据推理规制方面更为深入的标准是前述 DSMM 中的数据分析安全部分，基本上涵盖了数据推理规制的主要需求。DSMM 不仅要求设置专门负责的岗位和人员，还明确要求建立操作日志记录、数据分析安全审核、监控审计措施、合规评估等机制及相关技术工具，并要求记录与保存数据处理与分析过程中对个人信息、重要数据等敏感数据的操作行为，甚至明确规定了数据分析安

全审核的内容(数据源、数据分析需求、数据分析逻辑等)和监控审计措施的目标(确保实际进行的分析操作与分析结果使用与其声明的一致)。这一标准仍在继续修订完善之中，鉴于其对数据分析安全提出了相当系统的规制方案，未来将有可能成为指引数据推理规制的重要依据。

上述技术标准构成了规制单一数据库中数据推理活动的“软法”机制，尽管具体细节或未彻底成型，已能在防止不当数据推理方面承担稳定的规制功能。

(二)复杂数据来源的数据推理规制

真正棘手的难题是复杂数据来源的数据推理规制。这里的“复杂数据来源的数据推理”主要是指跨多个数据库的推理，以及部分关键推理步骤在数据库之外进行的数据推理活动，如此数据库内置的各种工具难以全然发现数据推理的真实目的和路线。理论上，结合不同数据来源、旨在发现敏感信息的数据推理是很难被彻底阻止的，只能尽可能在数据推理滥用风险与数据资源开发利用之间取得最优的平衡。这一平衡的具体确定，须视当前法律与监管政策框架中对某一类、某一级数据的安全施加何种保护力度。保护敏感数据的关键通常不是保护数据的主键(尤其是代理键),而是保护相关数据的敏感属性。因此，我们可以将数据安全与数据开发利用的价值平衡，视为需要保护的敏感属性与可利用的数据属性相结合的优化处理：如果在特定场景中，绝对强调数据安全的优先性，则此种优化可以被看作以绝对不泄露需要保护的敏感属性作为约束条件，追求可利用数据属性的最优化暴露(极端条件下需要完全禁止相关数据的共享、开放或流动);如果相关场景不要求数据安全的绝对优先性(例如基本不涉及核心数据、重要数据保护及个人敏感信息保护等),则此种优化可以被看作设置敏感属性的泄露风险阈值作为约束条件的情况下，追求可利用数据属性的最优化暴露。由此，对不同数据的推理规制之具体平衡，取决于不同数据在分类分级保护中的实际保护程度，许多细节仍需要等待完整的数据分类分级保护体系出台方可确定。

尽管尚无确切的分类分级规制方案，在上述基本原理的基础上，对于复杂数据来源的数据推理规制，若干法律机制也已显示出应用的必要性：

一是数据利用的目的限制。目的正当原则早已在大数据侦查等领域得到有力的提倡。[26]法律规范可以对数据推理提出专门性的要求，强调不得超出数据利用的正当目的范围进行数据推理；尤其是对于利用重要数据、核心数据等可能包含敏感信息的数据进行分析处理活动的，应当表明数据利用之目的，并接受相应的监督和审计，防止数据推理的滥用。

二是技术标准及相关的法律授权。类似于数据库系统管理，针对数据交易等数据处理活动及隐私计算、联邦学习等技术或业务，法律规范也可以授权建立相关技术标准，尽可能保证在“可用不可见”的前提下避免泄露原始数据或关键参数，防范数据推理对敏感信息的威胁。[27]

三是安全检测、评估、认证等专业第三方介入机制。在《网络安全法》《数据安全法》《个人信息保护法》中有大量此类机制，包括但不限于国家对数据安全检测评估、认证等服务的支持机制、重要数据安全风险评估机制、个人信息保护影响评估机制、网络安全风险评估机制等。数据推理风险规制是高度专业化且风险动态演化的领域，因此可靠的第三方服务必不可少，专业第三方的介入也更有利于检查技术标准是否在实践中得到贯彻实施。在现有法律规范的基础上，未来的配套立法应进一步明确第三方参与数据推理规制的具体要求。[28]

四是数据使用授权动态控制机制。在“原始数据不出域”的数据利用方式日益占据主导的前提下，如果数据利用越来越多地依赖开放数据接口和隐私计算处理，可以要求数据使用者在使用数据前声明数据分析需求，对各种主体利用数据的情况进行记录、审核和分析评判，并结合数据使用者自身可能掌握的数据类型及使用历史进行风险预测，在发现重大数据安全威胁或发生数据安全风险事件后，及时关闭或调整数据使用授权。

五是数据合规体系。法律可以要求达到一定体量的数据处理者建立数据安全合规体系，对包括数据分析在内的数据处理操作实行全过程记录并采取监控审计措施，确保数据推理始终在合法的数据分析目标范围内进行。

六是数据安全攻防演练。前文提及的“大数据安全靶场”即为可以开展数据安全攻防演练的平台。通过数据安全攻防演练，数据推理的潜在风险可以得到一定程度的暴露，进而为数据安全策略的调整和完善提供方向。

上述专门性法律机制似乎已经为数据推理的风险规制建立了一个完整的制度框架，可以大为提升数据推理规制的颗粒度和精确度。乍观之下，只要将这些机制堆砌起来，就似乎足以形成一整套规制数据推理的“组合拳”。然而，如果不认真剖析相关法律机制潜在的利弊得失，数据推理规制恐将陷于“多端寡要”的困境。

四、数据推理规制的路径选择

对于数据推理规制而言，既有的专门性法律机制尚处于探索阶段，难以完全断言何种机制将会取得可观的实效。然而，基于数据推理活动的特点，提前预判各种法律机制的应用潜力及潜在问题，对于建构完备的数据推理规制体系至关重要。在纷繁复杂的各种法律机制中，一个根本性的问题似乎甚少在学术讨论中被充分考虑：数据推理规制以及类似的数据安全治理机制，重心应当如何确定?这一问题的回答决定了数据推理规制的路径选择。

(一)第三方规制

经由专业第三方提供的风险检测、评估、认证、审计等服务进行规制，在新型风险治理的有关研究中日渐成为焦点，数据治理中也不乏提倡第三方规制的主张。[29]然而，对于数据推理规制而言，仔细分析若干代表性的第三方规制措施，或许会得出有所保留的结论。

数据安全认证或许是最早为法学界所关注的数据安全法制之一，也是数据安全领域典型的“第三方规制”路径。前述《数据安全法》等法律规范已有相关的原则性规定，属于国家支持开展的业务。虽然“数据安全认证已逐渐成为全球数据治理的重要手段”,但即便是数据安全认证的积极支持者，也承认“不同于传统认证，数据安全认证的对象主要是网络服务或数据产品，具有虚拟性和动态易变性的特征，认证难度更大”,继而主张“区分不同的认证领域与事项，结合数字科技易变的特点，根据具体情况合理确立认证标志的有效期”。[30]问题是，数据安全认证的对象并不是传统意义上的商品或服务，而是数据“产品”(本质上其实也是服务)和数据处理服务的安全状况,后者往往不能依赖单独分析,而需要结合数据处理者的整体管理制度、软硬件设施、技术路线和水准加以判断。然而,认证机构与腾讯、百度、微软、亚马逊这样的大型企业在技术储备和专业能力上往往相距甚远,即便部分员工本身可能有相关企业的工作经验，也会出现若干难以克服的问题：一是容易出现不正当的利益关联，导致认证的中立性缺失；二是企业间的数据处理活动在技术路线和软硬件设施方面可能相去甚远，依凭一个企业的任职经验难以评判另一个企业的数据安全水平；三是相关企业本身就规模庞大、部门众多，有复杂的业务流和专业分工，甚至在全球范围进行业务和组织的分割与整合，即便身处其中，亦很难掌握其数据安全的完整真实状况；四是数据技术及相关数学原理发展迅速，一旦脱离相关企业的高水平团队及平台资源支持，一段时间后很可能跟不上数据推理技术的前沿进展。因此，数据安全认证之类的单纯第三方规制手段很难在数据推理规制中担纲主角，只能起到有限范围内的辅助作用。

较之数据安全认证，数据安全攻防演练是一种不十分纯粹但亦包含了第三方规制思维的规制手段，治理者本身是组织者，多个专业第三方通过一定演练方案参与其中。理想情况下，数据安全攻防演练有助于组织多方专业力量，及时发现和准确评估数据推理方面的潜在威胁。然而，这也意味着多方专业人士有机会深入接触新的数据技术、应用和产品潜在的推理风险，在数据推理安全风险高度隐蔽、各方主体信息和能力高度不对称的前提下，组织者难以知悉参与者是否发现和报告了全部风险点，在有限的预算约束条件下，也不易找到满足“直接揭示机制”[31](direct revelation mechanism)标准的激励方案激励参与者报告全部风险点。不仅如此，对借助数据推理获取敏感信息有特殊需求的个人、组织或利益集团，也有可能设法通过主动参与数据安全攻防

演练获取相关信息。若为防止信任风险的发生而使用替代性的模型及数据，数据安全攻防演练的效果也会随之下降，模型及数据结构越失真，理论上攻防演练能发现的有用安全信息就越少。因此，数据安全攻防演练在专业能力问题之外，还需要着力解决可信任性与演练效果的平衡问题。较为成熟、完善的数据安全攻防演练组织方案成型后，此种机制才能成为规制数据推理安全的主要手段。

由上述两个例子不难发现，第三方规制在这样一种能力和信息高度不对称的局面中很难发挥主要作用，能力问题和信任问题都足以实质性地影响数据推理规制的成效。当然，这不排斥第三方规制的法律机制在能力结构较为均衡、信息和激励较为充分的某些场景中正常发挥作用，数据推理规制依然需要不断支持和完善专业第三方参与规制的法律机制。

(二)行政机关的直接规制

如果“第三方规制”难以在数据推理规制中担纲主角，那么将重心放在“第二方”(行政机关)[32]是否可行?回归行政机关直接规制的路径似乎也很有吸引力，但不易克服的难题亦如影随形。

首先，行政机关在数据推理规制中的直接信息来源可谓远不敷用。基于保护商业秘密、个人信息及规制成本等多方面的原因，大企业以及利用海量数据的其他组织很难向行政机关呈现数据处理活动的实时截面或报送全部数据；与此同时，行政机关也缺乏足够的人力分析海量数据处理活动中潜在的推理风险。如果没有相对方(被规制者)的有力配合，试图从零开始在EB以上级别的天量数据(其中许多还是加密后的数据)面前发现数据推理风险，不啻于“挟太山以超北海”。

其次，行政机关的技术力量往往不足以与大型企业及顶尖研究团队相抗衡。尽管网信、工信等行政机关中不乏专业人士，且还有相应的研究机构(如信息通信研究院)为之提供助力，但与全国乃至全球前列的平台企业所储备的技术力量相比，仍存在一定差距。不仅如此，如果采取直接规制的方式，规制者可谓“进攻方”,相对方是“防守方”。在数据推理活动可以存在较强信息不对称特性的情况下，如果规制者对平台企业缺乏持续的穿透性观察和必要的导引，即便前者组织全部执法力量“强攻”平台企业，只要后者预作准备，即有机会掩盖其实际上进行的数据推理活动。

再次，直接规制成本的高昂亦足以使规制者望而却步。且不论直接规制必然对监管人员提出大幅扩编要求，从海量数据中发现和消除推理风险所需的物力资源也是沉重的负担。要保证每一次调取数据的安全和维护巨大体量数据处理活动正常进行，耗费的公共财政资源可能是许多地方难以承担的。高昂的规制成本不仅可能引发来自比例原则和成本收益分析的质疑，甚至有可能使数字经济的巨大发展机遇变成沉重的社会负担，这是不可接受的结果。

直接规制极易引起的更深层问题还包括规制僵化、动力匮乏、规制俘获(regulatory capture)、监守自盗和规制过度等等。[33]若需再建立相应的监督制度精细入微地监督工作人员乃至监管机构自身是否利用这些数据进行不适当的推理活动，不仅监督成本会相当高昂，专业性及可行性方面亦无从保障。因此，监管者组织的数据安全风险评估、数据安全审查等带有直接规制性质的规制手段尽管仍有必要保留和发展，但这些“直接”规制手段往往也借助了监管机构和专业第三方的合力，且很难独力支撑数据推理规制的需求。

(三)元规制：针对数据分析安全的合规体系

对上述数种专门性法律机制之分析可以表明，数据推理的法律规制尽管需要借助专业第三方的力量，但其重心不可依赖所谓的第三方规制，而单纯依赖行政机关的直接规制亦不现实，关键是要让拥有强大数据推理能力的相对方积极参与规制过程，亦即通过“元规制”(后设规制)的方式使之主动建立数据处理的合规体系。元规制方式的优势在于承认实现规制目标的能力主要掌握在被规制者手中，并因势利导设计针对性的规制方案。[34]对于企业而言，企业合规的本质是一种自我监管(self-policing)机制，合规体系的建立意味着企业建立了商业行为规范、为员工确立了行为准则，并建立了合规风险防范体系、风险识别体系和

违规行为应对体系。[35]在数据治理领域，建立合规体系之意义更在于促进数据处理过程结构化和标准化，设置数据处理过程中清晰、精细的“问责点”，为企业内部的自我规制以及监管机构和专业第三方的“穿透式”监管奠定良好的基础。

针对数据推理活动，根据数据治理活动中已有相关标准及业务实践，合规体系的要点可以概括为“入库”“留痕”“控权”“审计”“设岗”五点。所谓“入库”，是要求数据处理者将数据汇入技术上安全、成熟的数据库，利用符合一定安全性标准的数据分析工具进行分析，便于数据处理者及监管机构运用各种内置的管理工具进行检查与监督。“留痕”，是要求保留数据访问和处理的操作日志，特别是对敏感数据的操作记录，一旦发生数据安全风险，可以进行回溯分析。“控权”，是精确控制访问权限，尽可能避免内部人员出于不正当动机滥用数据进行推理或分析活动，对于敏感数据，更要实行严格的权限控制，使数据安全风险能够精确对应到责任人。“审计”，是使数据处理活动处于监控审计之下，可以运用智能化的技术工具，及时发现异常的数据访问和处理活动，检查数据分析之目的是否与分析者声明的目的一致。“设岗”，是设置专门保护数据分析安全的工作人员，对异常的数据推理行为及时介入和调查。上述要点均已不同程度被现有的数据治理活动及技术标准吸纳，对于数据推理风险的控制而言亦可发挥相当有力的作用。

在前述五项要点的基础上，根据数据推理活动的规律，合规体系还可以进一步提出专门性的要求，例如对部分类型和级别的数据进行加密处理、审核数据分析需求、利用安全的数据分析工具等，并使每一点都对应不同层级的既有法律规范，进而为监管机构的监督检查提供充分、清晰的证明材料。其中，最有针对性的专门性机制是要求数据处理者内置推理风险发现工具(包括可对用户开放的风险警示工具)，并对发现的数据安全风险进行记录和报告。例如，针对基于线上社交网络的公开可用信息推断高度敏感个人属性的风险，已有多种量化数据属性推断风险的模型或工具，提醒用户保护其个人敏感属性免遭跨平台推断之侵害。[36]在未来，对于数据分析安全而言，合规体系需要更充分的风险监测技术支持。在确保“原始数据不出域”的前提下，对于多源敏感数据的跨企业聚合利用情形，监管者可以积极引导企业通过数据交易平台在安全的数据分析系统内进行聚合利用，并在系统内部配置必要的推理风险监测工具和推理通道阻塞措施。对于多元敏感数据的企业内部聚合利用，则应要求企业建立一体化的数据安全治理机制，并引导企业内置基于行为模型的风险预警工具，发现员工对敏感数据的操作偏离正常行为模型时，及时记录和发出预警信息，必要时直接限制访问权限，如果涉及重大公共利益，还应当同时自动直报监管机构。在此基础上，监管机构还可以针对不同数据处理者掌握的数据种类和内容进行敏感数据聚合的模拟推演，并基于推演结果要求掌握相关数据处理者对有风险的信道(包括侧信道)采取可追溯的保护措施，防止多源敏感数据的非法聚合。

当然，上述合规要点并非“一刀切”式地适用于所有类型、等级的数据及所有数据处理者，关键是在数据分类分级保护制度建立后，对不同数据源及应用场景配置不同程度的安全保护规则，通过有差别的合规体系要求数据处理者吸收或内化数据安全价值相对于企业自身发展利益的“外部性”效应。[37]

因此，完善数据推理的法律规制体系，当务之急是要求数据处理者基于不同数据处理场景的技术、业务与风险特点，建立数据分析安全方面的合规体系，并将细颗粒度的规制要求内置于合规体系乃至技术工具之中。这不仅是降低规制成本的需要，更是充分回应此领域技术发展不平衡、不确定、信息不对称挑战的需要。合规体系可以迫使具有较强技术能力和丰厚数据资源的数据处理者在保持日常业务基本正常运行的基础上固定关键的监督与审查要点，从而实际上大幅缩短从监管机构到规制对象之间的“规制距离”。监管机构则需要“及时划定制度边界，减少制度与政策的不可预期性，为企业合规提供更为精确的指引”，[38]并监督相对方切实遵行合规要求，在合规失灵时进行必要的干预乃至问责。与此同时，数据分析安全合规体系的建设与维护也可以更好地吸纳专业第三方的参与，通过合规体系建设、合规评估标准构建与监督检查过程的相对分离，以及合规标准与体系建设过程中不同被规制者与专业第三方的意见交锋，一定程度上弥补信任的缺失。数据推理风险发现工具的开发和问责点的埋设，还可以按需逐项逐点吸纳有特定专长的数据处理者及专业第三方参与，通过“化整为零”的方式一定程度上弥补能力的差距。由此，数据分析安全合规体系可以更充分地吸纳数据开发利用和数据推理风险认知的前沿发展，并根据数据推理的研究和实践不断完善。

数据分析安全合规体系的建设和完善，不仅可以帮助监管者更便利、集约、有效地实施监管，可以帮助数据处理者完善企业

的内部管理体系与工作流程，更可以帮助数据处理者及相关从业者系统和完整地学习数据推理方面的数据安全知识，进而促进整个行业乃至全社会达成数据安全治理的相关共识，为数据资源开发利用乃至数字经济的未来发展奠定精细化的制度基础。

五、结语

随着《数据安全法》和《个人信息保护法》的实施，数据安全治理体系的初步建立，数据交易和开发利用的范式逐渐明朗，同时持续催生对数据推理活动的规制需求。数据本身并不必然具有价值，数据的价值取决于能够从数据中获取何种信息，对数据推理的精准规制意味着对数据中隐含的信息流向在较细颗粒度层面的精确控制，此种控制还需要在技术发展不平衡、不确定、信息不对称的条件下实现，探索数据开发利用与数据推理风险预防之间的法治化平衡诚可谓寻幽入微。此一平衡端赖于各方主体在数据治理中形成的合力，而数据资源与技术力量均属雄厚的平台企业不仅应成为数据推理规制的重心，也应承担起向市场和社会展示数据安全利用范本的责任。

数字经济的滚滚洪流正在席卷这个时代。在未来，全球范围内的国家和社会对数据的依赖将进一步加深，人类社会数据的体量将从 ZB 向 YB 乃至更高级别迈进，数据开发利用的方式更将远超当代人的想象。尽善尽美而确凿不变的数据推理规制方案或许并不存在，数据推理规制的平衡点将根据数字经济发展和数据安全形势的演化而不断变动，但抓紧建立以数据分析安全合规为重心的规制体系并使之在法治轨道上运行，是当前数据安全治理体系建设的一项重任。

作者图片

苏宇

注释：

1 参见彭江:《北京加快培育数据交易市场》,《经济日报》2021 年 4 月 6 日,第 11 版。

2 参见闫林、闫硕:《粒计算与数据推理》,科学出版社,2019 年,第 3 页。

3 徐铮、陈恭亮、李建华:《基于推理树的 XML 推理控制研究》,《通信技术》2015 年第 2 期。

4 See P.Schwartz & D.Solove,“The PII Problem:Privacy and a New Concept of Personally Identifiable Information,”New York University Law Review,Vol.86,No.6,2011,p.1821.

5 参见晋涛:《刑法中个人信息 “识别性”的取舍》,《中国刑事法杂志》2019 年第 5 期。

6 参见齐英程:《我国个人信息匿名化规则的检视与替代选择》,《环球法律评论》2021 年第 3 期。

7 See T.Kenekar& A.Dani,“Privacy Preserving Data Mining on Unstructured Data,” in Sharvari Tamane,Vijender Solanki & Sharvari Tamane eds.,Privacy and Security Policies in Big Data,IGI Global,2017,p.173.

8 徐岩等:《基于推理通道的函数依赖推理控制》,《计算机仿真》2008 年第 1 期。

9 See J.Hale and S.Shenoi,“Catalytic Inference Analysis:Detecting Inference Threats due to Knowledge Discovery,”Proceedings.1997 IEEE Symposium on Security and Privacy (Cat.No.97CB36097),1997,p.188.

10 See Vicenç Torra,Data Privacy:Foundations,New Developments and the Big Data Challenge,Gewerbestrasse: Springer International Publishing,2017,p.67.

11 参见封皓君、段立、张碧莹:《面向知识图谱的知识推理综述》,《计算机系统应用》2021 年第 10 期。

12 参见杨庚、王周生:《联邦学习中的隐私保护研究进展》,《南京邮电大学学报》(自然科学版)2020 年第 5 期。

13 雷娜、顾险峰:《最优传输理论与计算》,高等教育出版社,2021 年,第 393 页。顾险峰教授对这一观点曾在不同场合进行专门的解释,此处的“算法”实际上是指算法模型。

14 参见王璐璐等:《机器学习训练数据集的成员推理综述》,《网络空间安全》2019 年第 10 期。

15 See Nicholas Carlini,Florian Tramèr,et,al.“Extracting Training Data from Large Language Models,” 30th USENIX Security Symposium (2021),p.2633.

16 参见李景海、唐明、黄诚轩:《基于侧信道与量化推理缺陷的模型逆向攻击》,《网络与信息安全学报》2021 年第 4 期。

17 See T.Orekondy,B.Schiele and M.Fritz,“Knockoff Nets:Stealing Functionality of Black-Box Models,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),2019,pp.4949-4958.

18 参见苏宇:《数据安全的双重向度及制度回应》,《中国社会科学报》2021 年 11 月 10 日,第 4 版。

19 参见李宗育等:《同态加密技术及其在云计算隐私保护中的应用》,《软件学报》2018 年第 7 期。

20 参见赵精武、周瑞珏:《隐私计算技术:数据流动与数据安全的协同保护规则构建》,《信息通信技术与政策》2021 年第 7 期。

21 参见苏宇:《数据安全的双重向度及制度回应》,《中国社会科学报》2021 年 11 月 10 日,第 4 版。

22 对于云计算平台等共享数据池而言,细粒度的访问控制是平衡数据安全与利用必须解决的问题。参见韩德志、吴帅、毕坤:《一种在云计算下的细粒度数据访问控制算法》,《华中科技大学学报》(自然科学版)2012 年增刊第 1 期。

23 See Jessica Staddon,“Dynamic Inference Control,” DMKD03:Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery,2003,p.94.

24 参见黄晓森、彭利宁、陈启买:《基于数据立方体的动态推理控制方法》,《计算机工程》2011 年第 17 期。

25 基本安全原则是一个数据项的安全类级别应支配所有施加影响于该数据项的安全类级别。防止推理的方法则包括“对数据重新分级”及“对约束重新分级”。

26 参见程雷:《大数据侦查的法律控制》,《中国社会科学》2018 年第 11 期。

27 参见赵精武、周瑞珏:《隐私计算技术:数据流动与数据安全的协同保护规则构建》,《信息通信技术与政策》2021 年第 7 期。

28 已有学者在算法规制领域提出了系统性的第三方介入治理的制度设计主张,可供数据推理规制的第三方介入机制建设作参考。参见郑智航:《人工智能算法的伦理危机与法律规制》,《法律科学》(西北政法大学学报)2021 年第 1 期。

29 例见翟志勇:《论数据信托:一种数据治理的新方案》,《东方法学》2021 年第 4 期;张继红:《大数据时代个人信息保护行业自律的困境与出路》,《财经法学》2018 年第 6 期等。

30 刘权:《数据安全认证:个人信息保护的第三方规制》,《法学评论》2022 年第 1 期。

31 直接揭示机制是机制设计理论的代表性成果之一,致力于通过一定激励机制(包括正向和负向的激励)迫使机制中的个体直接报告真实信息(如对某选项的真实意愿及利益相关程度)。See Dirk Bergemann and Juuso Välimäki,“Information Acquisition and Efficient Mechanism Design,”Econometrica,Vol.70,No.3,2002,pp.1007-1033.

32 对第三方规制语境中“第一方”“第二方”所指为何有不同的理解,但主流理解中“第一方”是指被规制的企业,“第二方”是指政府。关于不同指称对象的详细梳理参见刘亚平、游海疆:《“第三方规制”:现在与未来》,《宏观质量研究》2017 年第 4 期。

33 参见畠山武道:《行政介入的形态》,鲁鹏宇译,《当代法学》2012 年第 5 期;谭冰霖:《环境规制的反身法路向》,《中外法学》2016 年第 6 期;杜辉:《挫折与修正:风险预防之下环境规制改革的进路选择》,《现代法学》2015 年第 1 期等。

34 参见程莹:《元规制模式下的数据保护与算法规制》,《法律科学》(西北政法大学学报)2019 年第 4 期。

35 参见陈瑞华:《论企业合规的性质》,《浙江工商大学学报》2021 年第 1 期。

36 See H.Simo,H.Shulman,M.Schufrin,S.L.Reynolds and J.Kohlhammer,“PrivInferVis:Towards Enhancing Transparency over Attribute Inference in Online Social Networks,” IEEE INFOCOM 2021-IEEE Conference on Computer

Communications Workshops,2021,pp.1-2.

37 参见洪延青:《国家安全视野中的数据分类分级保护》,《中国法律评论》2021 年第 5 期。

38 张凌寒:《平台“穿透式监管”的理据及限度》,《法律科学》(西北政法大学学报)2022 年第 1 期。