

数据分类分级制度下的重要数据识别研究

杜聪艺

澳门科技大学 澳门 999078

摘要：在数据分类分级制度中重要数据处于承上启下的关键位置，为了保护数据安全，对重要数据进行识别至关重要。本文指出了重要数据本质上属于数据分级制度，厘清了重要数据与个人信息、国家秘密的关系，并基于《数据安全法 数据分类分级规则》梳理了重要数据的识别框架，即通过定性分析与定量分析相结合的方法进行综合判断。强调了行业主管部门与数据处理者在识别工作中的主体责任。本文旨在为构建重要数据识别与管理提供些许帮助，以促进数据安全流动。

关键词：重要数据识别；数据分类分级；定性定量分析

一、问题提出

随着数据时代的降临，数据从传统信息交换工具转变为能够促进全球贸易、技术创新的关键性生产要素。但与之相伴的是，基于对数据传播安全的考量，数据的保护和利用矛盾日益凸显。我国作为全球数字贸易大国，数据安全不仅关乎其自身的利用，还与国家安全、经济运行、社会稳定、个人和社会组织权益息息相关。在此背景下，数据需要多元化保护利用措施。然而，传统的数据安全保护制度对所有数据的保护都“整齐划一”并采用了一种重静态轻动态的数据安全保护策略，无法应对数据时代数据异质性特征和数字贸易时代数据流动常态化。对所有数据采用一套保护措施可能导致普通数据被过度保护，而与国家安全、社会公共利益有关的特殊数据没有得到应有保护。重静态轻动态的数据安全保护策略则会造成数据传播流动过程中的制度漏洞，尤其在面对数据跨境传输方面往往会束手无策。由此，为了落实数据保护，数据分类分级制度应运而生。数据分类分级制度根据数据自身的重要性和数据泄露后的危害性对不同数据进行分级，并在此基础上对不同内容形式、不同流动性的数据采取差异化管理。^[1]而本文讨论的对象重要数据正是数据分类分级制度的重要组成部分，对重要数据的保护亦是维护好数据安全的关键。进一步而言，重要数据在数据分类分级框架中处于核心枢纽位置，向上衔接国家核心数据，向下关联一般数据。对一般数据的规制思路以利用为主，保护为辅，所以对一般数据的保护只需符合数据安全的基本要求即可。而对国家核心数据的规制思路则应以保护为主，利用为辅，对国家核心数据的保护措施要强于一般数据和重要数据。至于重要数

据，其介乎一般数据和国家核心数据之间，处于一个特殊位置，对其的规制需要兼具保护和利用双重目的，既不能为了保护而完全限制重要数据流动，又不能为了利用而忽视保护或者减少保护措施。若能保护好重要数据，发挥其承上启下的作用，数据安全在一定意义上就能得到根本保障。为了配合数据分类分级制度的落地，规范重要数据识别，全国网络安全标准化技术委员会和国务院相继颁布了《数据安全法 数据分类分级规则》（以下简称“规则”）和《网络数据安全条例》（以下简称《条例》）。尽管近年来我国在立法层面高度重视重要数据的规范，但是在学理层面上缺少对重要数据的研究。以“重要数据”为关键词在知网上进行搜索，只有十篇左右的相关文献资料，且多集中于对立法原则和词意辨析的解读，没有对重要数据识别的深入研究。由此可见，研究重要数据识别具有一定价值。首先需要明确重要数据在数据安全体系中到底是分类制度还是分级制度？其次重要数据与国家秘密、个人信息等概念的关系如何？这是识别重要数据的前提。最后明确重要数据的内涵后，重要数据的识别标准到底是什么？

二、重要数据是分类制度还是分级制度？

重要数据的识别工作是数据分类分级制度落地执行的关键环节。《中华人民共和国数据安全法》（以下简称《数据安全法》）第21条作出明确规定，数据分类分级的标准是数据的“重要程度”和“危害程度”，其中重要程度是指数据本身的价值和作用，而“危害程度”是指数据泄露或破坏后带来的后果。从立法表述来看，立法者并没有区分数据分级和数据分类，认为分类和分级是并行维度。

正因为《数据安全法》没有明晰，所以我们在识别重要数据前需要解决重要数据到底是数据分类还是数据分级的问题。学界对这一问题也有争议，蓝蓝教授认为，重要数据属于数据分级制度，而非数据分类制度。而袁康教授认为重要数据在根本概念上是一个数据类别，但在具体管理和保护上，这个类别内部需要再进行数据分级。本文认为数据分类主要服务于数据的管理和使用，根据数据的内容主题、来源、业务领域等不同维度，将数据归入不同的类别，并且同一个数据可以根据不同维度被归入多个类别。例如，一份患者诊疗数据，从数据主体看是个人信息，从行业领域看是卫生健康数据，从描述对象看是业务数据。数据分级则是主要服务于数据的安全保护，根据数据一旦遭到安全事件，可能对国家、社会公共、个人和组织造成的危害程度，对数据进行等级划分。尽管《数据安全法》第21条并未细分数据分级和数据分类，但是立法者将该条款归置于数据安全制度章节，与数据分级的目的相同，整个条款从内容上来看也更像是数据分级的相关规定，由此本文推断重要数据制度应属于数据分级制度。这一判断在最新公布的《规则》与《条例》中亦能得到支持。

三、重要数据的定义

学界对重要数据定义的争议主要集中于对重要数据和个人信息、国家秘密关系的探讨。

（一）重要数据与个人信息

《规则》3.2条和《条例》62条都明确规定了重要数据的定义。《规则》中指出仅影响组织自身或公民个体的数据一般不作为重要数据，所以个人信息原则上不被纳入重要数据的范畴。这种排除性表述并非否定个人信息的价值，而是个人信息已形成独立的管理体系，为避免制度重叠与执行冲突，通常个人信息不适用重要数据的规制要求。但是经聚合分析形成的批量个人信息可能因为具备更高的安全风险而被认定为重要数据。

（二）重要数据与国家秘密

重要数据与国家秘密之间究竟有何差别？国家秘密和重要数据是两个独立的概念。即使是《规则》中提到的比重要数据更为重要的核心数据，其也不属于国家秘密。与重要数据相比，国家秘密更为敏感，但是重要数据在多数情况下同样不宜公开。从规制思路上来看，对国家秘密的保护主要在于严格管控国家秘密的知悉权限，而对重要数据的保护则需聚焦于防范数据泄露、杜绝数据篡改，同时保障数据自身的真实性。^[2]

四、重要数据的识别规则

“重要数据”是数据分类分级制度的核心概念，对其的识别与保护是关乎国家安全与公共利益的重大议题。重要数据识别，指在立法层面界定其核心内涵后，于实践应用场景中从海量数据资源里精准甄别筛选出重要数据，进而构建重要数据目录，并对其实施强于一般数据，弱于国家核心数据的保护和管理措施。在重要数据识别领域，迄今为止最主要的参考文件是《规则》。尽管《规则》只是一个推荐性标准，法律效力较低。但是《条例》吸纳了国家标准《规则》对“重要数据”的界定，可见《规则》对研究重要数据识别有一定的意义。通过梳理《规则》，可以总结出如何识别重要数据，即《规则》构建了一套以基本原则为识别前提，以定性分析和定量分析相结合的识别体系。

（一）重要数据识别的基本原则

第一项原则“科学实用”指出数据分类分级工作必须基于数据自身的客观属性和特征，确保分类过程逻辑清晰，分类结果科学实用，避免主观臆断。第二项原则“边界清晰”强调数据分类分级时，重要数据与核心数据、一般数据之间的定义和判定标准必须准确、无歧义，确保不同级别的数据间界限分明，以便实施差异化的安全保护措施，能有效避免因级别识别模糊导致的数据保护不足或数据保护过度问题。第三项“就高从严”原则揭示了当数据可能涉及多个影响对象或产生不同程度的影响时，应按照可能造成的最高危害来确定其级别。若影响程度存疑，则应从严认定。该原则旨在确保风险防范全覆盖，避免因就低评估导致数据保护不足，是数据安全审慎监管规制思路的集中体现。第四项“点面结合”原则，主要考虑到既要评估“点”上的单项数据安全风险，也要综合分析“面”上多领域、多群体、多区域的数据汇聚融合后可能产生的叠加效应与系统性风险，从而综合确定数据级别，确保数据在聚合场景下的全面保护。数据分级结果并非一成不变，数据的业务属性、重要性和发生安全事件后的危害性都会随着时间的变化而变化，因此需要定期审核分级结果与更新数据目录。故第五项原则“动态更新”指出数据分类分级工作完成后，当数据的业务属性、重要程度或所属环境发生变化时，须及时重新定级，确保分类分级结果始终真实反映数据的当前状态与所面临的潜在风险，实现持续有效的安全保护。

（二）构建定性分析和定量分析相结合的识别体系

定性分析是重要数据识别的核心逻辑，即“数据影响分析”，主要分为两步进行。第一步是识别潜在影响对象，数据影响对象指的是数据受损后可能危及的对象。《规则》中规定了6种影响对象，分别为国家安全、经济运行、社会秩序、公共利益等。其中，国家安全、经济运行、社会秩序、公共利益是重要数据识别的主要关切对象。附录E详细列举了各影响对象的具体考量因素。例如，判断是否影响国家安全时就需考虑政治安全、国土安全、科技安全、生物安全等因素。第二步是判定危害程度，即评估安全事件对上述对象可能造成的危害程度有多大。《规则》将危害程度分为“特别严重危害”、“严重危害”和“一般危害”。附录F提供了具体的参考示例。例如，对“经济运行”造成“严重危害”的情形包括“直接影响宏观经济运行”或“对行业领域或地区的经济发展造成严重危害”。

定量与特征分析主要是通过识别数据的某些关键要素来进行分级。《规则》第6.3条及附录C指出，数据的某些内在属性和特征是其可能产生重要影响的客观基础。在识别时，应重点考察领域、群体、区域、精度、规模、深度、覆盖度与重要性等分级要素。在考察领域时需要关注数据是否涉及关键信息基础设施、国家安全重点领域、重要民生领域等。考察群体时须注意数据是否覆盖大规模人口或特定敏感群体。在涉及区域数据时，须厘清数据是否涉及重要地理区域、未公开的关键场所等。数据的精度、规模、深度也不容忽视，在进行识别工作时，须考察数据是否达到一定的精确度（如高精度地理信息）、是否具有一定的数量规模（如海量人口健康数据），或是否为通过研究分析得到的深层次敏感信息（如产业供应链薄弱环节、经济运行内在规律）。最后需要考察数据的覆盖度与重要性，如数据在特定领域、群

体或区域中的覆盖广度和密度，以及其在国家经济社会发展中的战略价值。具备上述一个或多个要素的数据，其成为重要数据的可能性会显著增加。

（三）参照具体指南进行综合判定

《规则》附录G《重要数据识别指南》是极具操作性的规范性文件。它列举了多达十余类可被识别为重要数据的具体情形，形成了一份“负面清单”性质的识别指引。这份指南将抽象的原则转化为相对具体的场景，为各行业各部门制定重要数据目录提供了帮助。识别工作应是在上述定性分析与定量特征筛查的基础上，结合附录G的具体指引，进行的一项综合性判断。

（四）落实数据识别的主体责任

学界对于是否需要从行业分类角度制定重要数据识别规则一直存在分歧。在国家层面进行重要数据的行业分类或许会削弱行业的自主决策权，且难以与行业发展的现状相契合。重要数据之所以重要并非因其属于某特定行业，而是重要数据影响到了国家安全。若完全摒弃“分类”理念，重要数据的定义会过于宽泛，进而造成相关标准缺乏实操性，给实际管理工作带来诸多不便，所以这是两难问题。但从目前的立法来看，我国立法者把制定重要数据目录的任务下放到各行业和各领域，《规则》规定：在行业领域层面，行业领域主管（监管）部门负有首要责任，需根据《规则》和国家要求，制定本行业本领域的数据分类分级标准规范，其中必须明确本行业本领域重要数据识别细则。在数据处理者层面，数据处理者（企业、组织等）则需在遵循国家和行业标准的基础上，制定内部细则，开展数据资产梳理，实施数据分类分级，并最终审核上报重要数据目录，使重要数据识别工作在基层的落地。

参考文献：

[1]袁康,鄢浩宇.数据分类分级保护的逻辑厘定与制度构建——以重要数据识别和管控为中心[J].中国科技论

坛,2022,(07):167-177.

[2]景杰.重要数据跨境流动法律问题研究[D].重庆:西南政法大学,2022.

作者简介：杜聪艺（2002.5.24—），性别：女，民族：汉，籍贯：湖南省石门县，学历：本科，单位：澳门科技大学，从事工作或研究方向：宪法与行政法。