

论公共数据的认定标准与类型体系

沈 斌

【摘 要】公共数据的概念范畴作为公共数据治理的基石,并未达至立法统一和学理共识。无论是当前立法采行的以数据主体要素公共性为核心的归属标准,还是学界提出的以数据内容公共利用价值为内核的用途标准,皆有所短。基于公共数据之事物本质与价值特征,公共数据的范畴勘定应当坚持“主体+内容”双重公共性的标准,即只有数源主体和数据内容本身皆表征公共利益的数据才能作为公共数据。并负担相应的数据处理义务。主体要素的公共性以公共资金保障或政府授权经营为表征,内容要素的公共性则因作为公共产品供给活动的副产品而具备。根据主体要素与内容要素的公共性差异,公共数据呈现为公共性渐次降低的政务数据、公共非营利主体数据、公共营利主体数据和授权经营主体数据四类。

【关键词】公共数据;公共利益;公共产品;认定标准;双重公共性

【作者简介】沈斌,华中师范大学法学院讲师(武汉 430079)。

【原文出处】《行政法学研究》(京),2023.4.64~76

【基金项目】2020年度国家社科基金青年项目“数字政府建设自动化行政的法律回应研究”(项目编号:20CFX021)。

数字经济时代,数据取得生产要素的价值地位,成为经济社会发展的基础资源。公共数据作为规模最为庞大的数据类型和最有希望坚持开放性的数据资源更是日渐受到重视。然而,作为一个技术型、工具性概念,公共数据并不存在固有的确定内涵与外延。就立法来看,中央立法尚未明确界定作为新生事物的公共数据的概念,而地方立法中的公共数据规范概念则莫衷一是,甚至是大相径庭。公共数据内涵外延的立法分殊不仅导致法律系统内部概念的不协调,阻滞数据资源在横向区域间的流通融合,而且各层级政府界定的公共数据外延宽窄不一,还可能造成政府公共数据治理的混乱,影响公共数据共享开放和开发利用法秩序的合理建构。

毫无疑问,公共数据是具有公共性的数据。然而,私人主体控制的数据因其利用的非竞争性、非排他性和处理后果的强外部性,同样存在高度的公共性,^①公共数据与私人数据的法理边界较为模糊,无法藉由功利主义福利最大化的逻辑予以区分。正因

如此,学界在论述公共数据的概念范畴时存在“用途标准”与“归属标准”的纷争。公共数据的认定标准是合理框定公共数据概念范畴的逻辑基础,决定着国家公共数据资源体系建设的科学性和公共数据要素价值释放的充分性,构成公共数据治理这一鲜明时代命题的基础要义。公共数据价值利用和治理立法的双重需求,使得如何在法律上设置公共数据的概念范畴及其认定标准成为数字经济时代最为基础且富含智识挑战的学术议题之一。

基于此,本文拟从现有立法实践和学理研究出发,尝试提出一套契合公共数据属性特征和数据价值实现规律、融贯“数据归属”与“数据价值”要素的公共数据认定标准,为公共数据规范概念的科学设计提供学理指引,并尝试通过类型化的方式呈现公共数据的类型体系。

一、公共数据认定标准的立法与学理检视

追根溯源,公共数据之语词滥觞于地方立法实践,首先作为规范概念而出现。随着数据价值的深

度挖掘和呈现,“数据”逐步完成对“信息”的超越和替代,政府直接治理的数据范围亦不再局限于政务数据。《浙江省公共数据和电子政务管理办法》首创公共数据概念后,引领了地方公共数据立法的潮流,“公共数据”逐渐替代“政务信息”和“政务数据”成为地方相关立法的主概念,并引发学界对其内涵外延的探讨。

(一)公共数据认定标准的立法殊同

纵观既有立法,公共数据规范概念的立法表述在语言逻辑方面具有趋同性,大致可将其概括为,“公共数据是主体要素+内容要素+行为要素的数据”的规范逻辑。主体要素表征产生公共数据的公共管理和服务机构,内容要素指称产生公共数据的履行公共管理职责或提供公共服务活动,行为要素则指称产生公共数据的制作或获取方式。梳理公共数据三要素与其外延范围的逻辑勾连,不难发现主体要素对公共数据的外延范围具有决定性意义。具言之,就现有公共数据概念规范来看,内容要素通常由主体要素所决定,履行公共管理职责或提供公共服务是数源主体的法定职责。虽然数源主体也存在履行公共管理职责或提供公共服务以外的活动内容并产生数据,但现行立法对此基本未予考量,除极个别地方立法外,公共数据的外延范围仍然由概念规范中的主体要素所决定。行为要素则具有中立性,制作或获取数据仅仅描述的是主体要素产生公共数据的行为方式,不会对公共数据的外延范围造成实质性改变。整体来看,地方立法勘定的公共数据概念范畴主要采用了表征数据归属的主体标准,而表征数据价值的内容标准,尤其是表征数据产生方式的行为标准,并未被立法所采纳和重视。在此意义上,数据虽然不同于传统意义上的“物”,但地方立法以制作或获取数据的主体为标准界定公共数据概念范畴的立法逻辑,仍内在承袭了我国物权立法在界定公共财产范围时采纳的“归属标准”。^②

然而,各地方立法虽然都采纳归属标准主导的相同规范逻辑,但不同立法文本中公共数据规范概念涵摄的外延范围却各不相同,甚至是大相径庭。根据公共数据主体要素,即公共管理和服务机构外延范围的差别,可将各地方立法规定的公共数据规

范概念分为广义、中义和狭义三个层次。细言之,公共数据的主体要素分为公共管理主体和公共服务主体两者。广义的公共数据规范概念涵摄的主体要素较为宽泛,涵盖全部地方公共数据立法出现的公共性主体。在公共管理主体方面,不仅包括行政机关和人大、司法、监察等其他国家机关,而且也包括具有管理公共事务职能的事业单位和其他组织;在公共服务主体方面,不仅涵摄供水、供气、供电、公共交通等公共服务运营单位,而且也涵摄教育、卫生健康和社会福利等公共服务组织。广义的公共数据规范概念以《深圳经济特区数据条例》为代表,根据其规定,作为公共数据主体要素的公共管理和服务机构为“国家机关、事业单位和其他依法管理公共事务的组织,以及提供教育、卫生健康、社会福利、供水、供电、供气、环境保护、公共交通和其他公共服务的组织”。中义的公共数据规范概念涵摄的主体要素相对较窄,在地方立法中占据主导地位。或在公共管理主体方面将行政机关之外的其他国家机关排除在外,比如,根据《广东省公共数据管理办法》的规定,作为公共数据主体要素的公共管理和服务机构为“行政机关、具有公共事务管理和公共服务职能的组织”;或在公共服务主体方面仅涵摄部分主体类型,比如,根据《浙江省公共数据条例》的规定,作为公共数据主体要素的公共管理和服务机构为“国家机关、法律法规规章授权的具有管理公共事务职能的组织以及供水、供电、供气、公共交通等公共服务运营单位”,据此,教育科研、医疗卫生、社会福利等公共服务主体被排除在外。狭义的公共数据规范概念涵摄的主体要素则仅限于行政机关和法律、法规授权的具有管理公共事务职能的组织,而不包括公共服务主体,基本等同于政府信息公开适用的主体范围。此类公共数据规范概念在地方立法中较为少见,目前仅见于《吉林省促进大数据发展应用条例》。根据其规定,公共数据的主体要素仅为“行政机关以及具有公共事务管理职能的组织”。

(二)公共数据认定标准的学理殊

随着公共数据价值的日益彰显和地方公共数据立法的持续推进,学界关于公共数据的研究亦逐渐兴起。而公共数据作为技术型、工具性概念的本质

属性和地方立法中概念规范的莫衷一是,造就了公共数据概念范畴,尤其是认定标准的学术议题。整体观之,公共数据认定标准的现有研究大致可划归为遵循公共数据立法逻辑的归属标准和强调数据价值维度的用途标准两者。

归属标准的公共数据研究范式通常不将公共数据的概念范畴作为单独学术议题,即不对公共数据的认定标准进行过多论述,而是直接引用特定地方立法文本中的规范概念描述公共数据的内涵外延,或是套用现有立法逻辑描述公共数据的范畴,并在此基础上展开公共数据归属、开放等相关研究。例如,有学者直接引用《浙江省公共数据开放与安全管理暂行办法》中的概念规范,将公共数据描述为“各级行政机关以及具有公共管理和公共服务职能的事业单位在依法履行职责过程中获得的各类数据资源”^③。无独有偶,也有学者直接引用《上海市公共数据和一网通办管理办法》规定的概念,将公共数据定义为“各级行政机关以及履行公共管理和公共服务职能的事业单位在依法履职过程中,采集和产生的各类数据资源”^④。另有学者在梳理上海市、北京市和成都市公共数据立法文本中的概念规范后,提出“公共数据是公权力机关依法履职过程中采集并以一定形式记录保存的各类数据资源”^⑤。由此不难发现,公共数据认定的归属标准是立法逻辑在学理层面的延展和继续,更加强调主体要素对公共数据内涵外延的勘定价值。

用途标准的公共数据研究范式则通常将公共数据的概念范畴作为单独学术议题予以阐述,基于数据资源最大化开发利用的价值导向将数据资源的使用特征,即“应由社会公众基于公用目的而共同使用”,作为厘定公共数据概念范畴的核心标准。据此,数据承载内容信息的使用价值居于决定性地位,数据主体要素不再限于公共机构,行业协会、科研机构、社会组织和平台企业亦应涵摄于其中。比如,有学者提出,归属标准锚定的公共数据资源范畴并不契合其本质特征,不仅外延失之过窄,而且也存在漠视使用者利益和强化所有人专断的潜在风险。为此,应当以契合公共数据使用特征的“用途标准”取代“归属标准”,确保公众平等享有对公共数据资源

的使用权。^⑥同时,有学者虽然并未明确论述公共数据认定的用途标准,但明确提出,作为公共数据内核本质的公共性并非指数据主体的公私属性,而是指数据承载的信息内容是否关涉公共利益,并将国家机关及其他公共机构以外的社会主体所控制的与公共利益密切相关的数据纳入公共数据的学理范围。^⑦可以说,该论本质上也坚持基于数据内容的公共数据认定标准。此外,部分学者直接将公共领域或空间中产生的涉及公共利益的数据纳入公共数据范围,如商业性平台上“由公民参与网络活动而产生”的具有公共性的数据,本质上也是将数据用途作为公共数据的认定标准。总而言之,公共数据认定的用途标准更加强调数据内容要素的公共性,凡是数据内容具有公共性使用价值的的数据资源都应当纳入公共数据的范畴,至于数据主体要素是否具有公共性则在所不问。

(三)现有公共数据认定标准的缺陷

归属标准与用途标准在勘定公共数据范畴方面各有所短。就归属标准而言,根据主体要素的公共属性勘定的公共数据概念范畴虽然更加契合公共数据治理实践,便于数据共享、开放和安全管理等公共数据承载的数据处理义务的落实。但同时必须承认的是,归属标准实际上并非一个逻辑完整的公共数据认定标准,将归属于特定主体的数据认定为公共数据本质上是对数据属性的判断转换为对数据主体要素属性的判断,何种类型的数据主体要素具有公共性实际上仍然是一个有待解决的问题。对此,归属标准并未给出明确的答案。也正因如此,当前地方公共数据立法虽然都采纳归属标准设计公共数据概念规范,但在公共数据外延范围方面却相去甚远,存在广义、中义和狭义的公共数据概念殊。不唯如此,正如用途标准论者对归属标准的批判,归属标准勘定的公共数据范围并不周延,不仅将公共管理和服务机构控制的全部数据资源均作为公共数据资源存在合理性质疑,而且一些具有内在公共属性的数据资源因主体要素公共性阙如而无法进入公共数据范畴也留有遗憾。^⑧

就用途标准而言,尽管根据数据内容的公共利用价值界定公共数据能够将更大范围内的数据资源

纳入公共数据范畴,有利于借助公共数据共享开放机制释放数据资源的内在价值。但相较可能带来的增益,用途标准作为一种泛化的公共数据认定标准,存在的固有缺陷和消极影响可能会更多。首先,正如前文所述,数据因其利用的非排他性和处理活动的强外部性,本身即存在高度公共性,因此公共数据与私人数据的边界可能无法单纯藉由数据内容的公共利用价值予以厘定。换言之,用途标准无法为公共数据提供清晰明确的范畴,公共数据将作为弹性概念处于不确定状态,从而背离法律确定性之现代法治精神,不仅无法为数据生产者提供确定性激励,而且容易滋生规则制定者的专断和擅权。其次,根据数据内容的公共利用价值界定公共数据范畴,固然整体上拓展了公共数据的范围,契合数字经济时代数据资源最大化开发利用的价值导向。然而,就数据资源持续生产利用的角度来看,将非公共性主体制作或获取的数据纳入公共数据范畴并使之负担共享、开放等强制性法律义务,不仅现实阻力极大,而且将极大地消解上述主体付出“额头汗水”生产数据的积极性,在根本上无益于以数据资源为关键生产要素的数字经济的可持续发展。最后,公共数据外延范围的扩展还将造成更为棘手的治理问题。申言之,公共数据外延范围的拓展,使得基于公共数据这一概念分析工具形成的数据资源池在规模日渐庞大的同时,各公共数据集之间的同质性也日渐式微,日渐凸显的异质性则使得基于单一概念设计具有普适性的公共数据治理规则愈来愈困难。用途标准框定的外延更为宽泛的公共数据概念将进一步加剧数据异质性带来的治理难题。

二、公共数据认定标准的省思与改进

无论是当前立法采行的以数据主体要素公共性为核心的归属标准,还是学界提出的以数据内容公共利用价值为内核的用途标准,勘定公共数据概念范畴皆存在瑕疵与缺陷。公共数据认定标准亟待在省思公共数据之事物本质与价值特征的基础上予以改进和完善。

(一)公共数据认定标准的省思

作为一个技术型、工具性概念,公共数据并不存在固有的确定内涵与外延。然而,“生活之情势,尽

管或多或少会有所发展,但他总是自有其尺度和章法。事物所内含的章法即可以被称作事物之本质”^⑩。公共数据之事物本质体现为公共性要素资源的属性,其认定标准亦应循此章法而建构。循此章法脉络并结合归属标准与用途标准的缺陷,公共数据认定标准的改进应当明确如下三个层层递进的前置性问题。

其一,公共数据作为以公共利益为建制基础的概念,其与公共利益的逻辑关系应当是表征公共利益,而非关涉公共利益。毫无疑问,公共利益作为公共数据的建制基础和根本正当性来源,亦构成准确界定公共数据范畴的前提和基础。尽管公共利益本身作为不确定性概念并不存在统一认知的具体内涵,^⑪但这并非意味着公共利益在具体情境中无法度量。从功利主义的视角来看,公共利益应当是个人利益总和的最大化,对应着效率优先原则。^⑫循此进路,公共数据应当是作为公共数据治理更能实现数据价值,即产生最大效益的数据资源。然而,数据因其利用的非排他性和处理活动的强外部性,本身即与公共利益存在天然的紧密关系,亦即全部数据都存在纳入公共数据范畴的动因。这显然不符合数据治理的基础逻辑。在此意义上,功利主义社会福利最大化的视角无法适用于可以零边际成本无限复制使用的数据及其价值,从而无法藉此区分公共数据和私人数据。^⑬归根结底,功利主义的公共数据认定进路,本质上将关涉公共利益作为公共数据的正当性来源,但却因数据的特殊属性而丧失合理性,这也正是用途标准诸多缺陷的根源所在。事实上,公共利益还存在另一种度量方式,即“公共主体使用”,据此,唯有代表公共利益的主体方才符合公共利益。^⑭追根溯源,此种公共利益度量方式滥觞于《美国宪法第五修正案》中的征收条款,其规定“私人财产非因公用目的、没有正当补偿不得征收”。为控制行政权的滥用,早期的美国司法实践对“公共使用”采取了较为严苛的限制性解释,通过诸多“判例”^⑮明示“不直接服务于公共利益的私人团体不能满足公用的要求”^⑯。理性而论,上述公共利益度量方式虽因在大多数情形下过于严苛而逐渐被美国司法实践所摒弃,^⑰但却能够为公共数据的范畴勘定提供恰如其分

的新进路,即将度量公共利益的落脚点置于业已存在的公共性事物,而非从头进行功利主义的利益衡量。一言以蔽之,在关乎公共利益之功利主义进路失效的语境下,可以根据表征公共利益的限制性进路衡量数据的公共利益,进而勘定公共数据的范围。如此,不仅可以清晰呈现公共数据与私人数据的界限,而且亦能同时吸纳归属标准和用途标准分别强调数据主体要素和内容要素的合理性。并回避其消极影响。

其二,基于表征公共利益的限制性进路勘定公共数据范畴,核心在于判断数据主体要素和内容要素内含的公共性。主体要素的公共性表征相对较为清晰,公共主体与非公共主体的分野较为显见,通常无需借助于其他理论工具。然而,内容要素的公共性判断则较为模糊,处理不慎则可能落入“用途标准”的路径依赖,导致公共数据与私人数据界限丧失。本质而论,公共数据作为公共品,^⑩或可援引传统公共产品理论作为判断内容要素具有公共性的有效理论工具。根据公共产品理论,公共产品是为了满足社会公共需要而由政府直接组织生产、供给或授权部分社会主体生产经营的产品,包括纯公共产品和准公共产品两类,前者如国防外交、生态环境、公共安全等,后者如教育、医疗、交通设施等。无论何者都是业已存在的表征公共利益的事物。而数据作为特定主体行为活动的记录,是该行为活动产生的副产品。^⑪推而论之,特定主体提供公共产品的行为活动产生的数据自然带有公共利益属性,记录公共产品供给活动的数据可作为内容要素具有公共性的重要证成。在此意义上,公共产品就成为判断数据内容要素公共性的相对清晰且具有可操性的标准工具。

其三,公共数据认定标准的合理性最终取决于其涵摄的数据资源负担公共数据应当负担的数据共享、开放和管理等义务是否具有法理正当性与现实可行性。亦即,公共数据相较私人数据应当负担的义务是公共数据范畴勘定时必须考量的重要因素。某一数据一旦被认定为公共数据就将背负诸多法律义务和责任。^⑫可以说,公共数据作为概念分析工具的价值并非体现在通过厘清数据权属划定各方主体

的利益边界,而是表现为藉由公共数据作为公共性资源应当负担的数据共享、开放和管理等义务,推动数据资源的价值最大化实现。归根结底,公共数据因负担义务而产生的不利益最终需要由数据主体来承担,特定数据主体承担数据处理义务的合理性构成公共数据与私人数据界分的重要依据。如果特定数据资源在根本上不具备承担上述数据处理义务的能力,即负担公共数据处理义务而产生的不利益缺乏正当性与合理性,则强行将其纳入公共数据范畴将毫无意义,徒增利益协调和制度运行成本。职是之故,公共数据认定标准的具体建构必须嵌入数据负担义务能力的考量,将其作为认定公共数据的辅助性工具,尤其是当特定数据资源的主体要素和内容要素的公共性模糊时,可以根据该数据资源负担共享、开放或管理等义务的合理性反向推导其公共性,进而判断是否适宜将其纳入公共数据范畴。

(二)公共数据认定标准的改进

基于上述理论逻辑,公共数据范畴的界定应当坚持“主体+内容”的复合型认定标准,即只有当特定数据资源的主体要素与内容要素都具备一定程度的公共性时,方才应当纳入公共数据范畴。主体标准与内容标准相辅相成、相互印证、彼此限制,共同将公共数据的范畴框定在一个合理的范围内。

一方面,公共数据应当符合主体要素公共性的标准。主体要素的公共性是其负担公共数据处理义务之不利益的基本正当性证成,即只有公共性主体方才具有负担强制性的公共数据处理义务的能力。首先应明确的是,数据主体要素具体指代的应当是数源主体而非数据控制主体。由于数据资源的可复制性和使用的非排他性,数据主体要素事实上并未指代具体明确,存在数源主体和数据控制主体等数个与数据关系密切且外延相互交叉的主体选项。相较数据控制主体因数据的可复制性和非独占性而可能存在复数性,数源主体作为制作或获取数据的主体,不仅为数据生产付出了额头汗水,与数据权益关系更为紧密,而且具有唯一性,能够避免因复数主体性质各异造成的数据性质难以确定的弊端,从而更适宜作为公共数据性质认定的主体要素基准。

在此基础上,为避免当前数据归属标准因逻辑

不完整而造成的公共数据外延范围的不确定性,应当进一步明确数据主体要素公共性的具体表征问题。法理而言,数据主体要素的公共性可以根据数源主体履行职能所需资金来源判断。公共财政资金和社会捐赠资金作为带有浓厚公共利益属性的资金,或可作为表征数源主体公共性的基本指标。申言之,公共财政着眼于公共利益,履行公共职能,^③先天就是公共利益的化身,^④依靠公共财政资金履行职能的主体具有当然的公共属性。同时,社会捐赠资金属于社会公益的范畴,构成需受国家审计监督或主管部门财务监督的重要公共性资金,^⑤依靠社会捐赠资金维持组织运转的主体也是公共产品提供者,也具有当然的公共属性。职能履行主要依靠公共财政资金或社会捐赠资金的数源主体当然符合公共数据主体要素公共性的标准特征。除此之外,政府授权的非依赖公共财政资金或社会捐赠资金从事公共产品生产经营活动的市场主体,亦应当纳入公共数据主体要素的考量范围,理由如下:其一,就公共产品供给的角度而言,虽然公共财政资金和社会捐赠资金支持的公共性主体是公共产品的基本供给者,但基于经济效率之考量,公共产品,尤其是准公共产品也可能在国家干预的基础上由市场主体予以供给。现实来看,教育、医疗、大型基础设施等准公共产品领域存在大量的市场主体经营者。此类通过政府特别授权取得准公共产品生产经营活动资格的市场主体,虽然不直接依赖公共性资金实施公共服务,但本质上是替代政府向社会供给公共产品,其生产经营活动虽然具有营利性目的,但需受到政府的诸多限制和干预,也具有相当程度的主体公共性,从而能够作为公共数据的数源主体。其二,就数源主体义务负担能力的角度而言,政府授权主体通过政府授权不仅获得生产经营特定公共产品的资格,而且亦攫取了一部分公共治理权力,承担了一部分公共治理责任。^⑥在此意义上,政府授权主体是以治理者的角色参与公共产品的供给,并产生数据,在享受该角色带来的经济利益的同时也具有当然的负担数据处理义务的责任能力。其三,就数据资源价值开发利用的角度而言,公共数据是能够坚持开放性的数据资源,可以藉由其负担的共享、开放等强制性义务

最大程度地释放数据资源蕴含的价值。因此,政府授权主体制作或获取的数据纳入公共数据范畴也契合数字经济时代数据资源价值最大化开放利用的价值定位。

另一方面,公共数据也应当符合内容要素公共性的标准特征。公共数据应当是内容要素表征公共利益的数据。前已述及,公共产品作为满足社会共同需要的产物,^⑦内嵌社会公共利益的意蕴,能够作为数据内容具有公共性的基本表征。藉由公共产品的理论工具,可以为特定数据之内容要素是否具有公共性提供相对清晰的判断标准。换言之,公共数据应当是记录公共产品供给活动的数据,属于公共产品供给活动产生的副产品。由此,公共产品性质特征和范围边界的判断就成为公共数据内容要素公共性认定的前提和基础。而公共产品理论已经是一个较为成熟的理论体系,公共产品也具有较为显见的非排他性和非竞争性等识别特征,公共数据内容要素的公共性因公共产品轮廓的清晰显见而在绝大多数情形下都较容易判断。然而,必须同时注意到,公共产品与非公共产品的边界仍然存在一定程度的模糊性。^⑧纯公共产品领域的边界相对清晰,而准公共产品领域的边界则较为模糊。因此,公共数据内容要素的公共性就难免因公共产品的模糊性而出现难以判断的情形。此时,可以借助主体标准与内容标准相互印证的逻辑机理,藉由数据主体要素的公共性判断内容要素的公共性。申言之,公共产品是市场失灵的典型场域,^⑨因此需要由公共财政资金或社会捐赠资金维持运行的公共主体或者经政府授权的特定市场主体经营。推而论之,特定产品是否由上述公共性主体生产经营可以作为其是否为公共产品的在先验证,亦即只有上述公共性主体提供的产品方才可能构成规范意义上的公共产品,而尚未由上述公共性主体垄断生产经营的产品则可以暂不视为公共产品。如此,公共产品与非公共产品的边界会相对清晰,公共数据的内涵外延亦相对清晰可见。

深究之,内容标准的主要价值在于构成主体标准的重要限制,剥离主体标准框定的不符合公共数据之事物本质的数据资源。换言之,并非全部公共

性主体制作或获取数据皆属于公共数据的范畴,其制作或获取的部分数据将因不符合内容要素的公共性而被排除于公共数据范畴。《广东省公共数据管理办法》第2条第2款将公共企业、事业单位和社会团体实施公共服务以外产生的数据排除于公共数据范围,正是基于上述考量而作出的合理规定。

综上,基于“主体+内容”之双重公共性标准,可将公共数据概念规范性地表述为,职能履行受公共财政资金或社会捐赠资金支持的公共性主体及政府授权生产经营准公共产品的市场主体,因履行公共管理职责或提供社会公共服务而制作或获取的数据。在此基础上,公共数据内涵外延的认知应当明确如下三方面的内容:首先,公共数据概念范畴的描述并不考虑数据承载的个人隐私、商业秘密和国家安全等问题,换言之,某一数据并不会因为涉及个人隐私、商业秘密和国家安全等而丧失公共数据的性质。事实上,上述情形并不会导致数据主体要素和内容要素公共性的失却,因此也不会导致数据义务能力的丧失,而仅可能因高位阶法益保护之考量而产生暂时阻却数据资源公共性使用之效果,即构成数据义务实现的阻却事由。比如,内容涉及个人隐私的公共数据按规定一般列入有条件开放或不予开放的数据范围,而一旦数据经过脱敏处理或引进“可用不可见”等数据处理技术,阻却数据开放的事由消失,该公共数据就应当予以开放。其次,公共数据概念范畴的勘定也不考虑数据资源现实利用价值大小的问题,某一数据资源只要符合主体要素和内容要素的公共性标准就应当纳入公共数据范畴,既不能因某一数据价值潜力较小而将其排除于公共数据范围,也不能因特定数据资源价值潜力较大而将其纳入公共数据范围。据此,数字平台企业等非公共性主体制作或获取的数据,即使公共利用价值再高也不属于公共数据的范畴,此乃坚持公共数据与私人数据之界限和确保数据治理具备确定性激励的必然要求。再次,公共数据范畴勘定坚持主体要素与内容要素双重公共性标准造成的部分高价值数据资源的“漏网”,并不会造成明显的数据资源价值开发利用的低效率。以数据开放为典型表征的公共数据治理本质上是数字经济时代政府配置资源的最新表现

形式。相较市场机制,政府的资源配置功能因数据价值特征的特殊性确实某些场景具有优势。然而,这绝非意味着市场机制配置数据要素资源处于完全失灵状态,也并非意味着市场机制在数据资源配置场域之决定性作用的丧失。事实上,随着数据要素市场建设的逐步推进,市场在数据资源配置方面仍将发挥决定性作用,未纳入公共数据范畴的数据资源可以藉由数据互换、数据交易等市场机制释放其价值。不惟如是,公共数据与私人数据之间也不存在绝对化的壁垒,私人数据可以通过行政征收、政府购买或无偿捐赠等私人数据公共化方式转化为公共数据。^④藉此,不仅政府治理对特定数据的使用需求得以满足,而且数据价值公共利用的目的亦可以在较大程度上实现。

三、公共数据的类型体系

前文论述可较为清晰地描述公共数据的内涵外延,而为更直观地感知公共数据的概念范畴,下文将尝试通过类型化的方式呈现公共数据的外延范围。作为一个规模庞大的数据资源池,公共数据涵摄的数据范围仍然非常广泛。公共数据虽然共同具备主体要素和内容要素公共性,但事实上,其内部各数据集之间也呈现出一定程度的异质性和层次性。具体而言,一方面,公共数据的数源主体存在公共性程度的差异,既包括职能履行完全受公共财政资金保障的数源主体,如履行公共管理职责的行政机关;也包括职能履行部分受公共资金保障的数源主体,如提供公共服务的企事业单位;还包括职能履行完全依赖社会资本的数源主体,如政府授权经营的公共服务主体等。履行职能所需资金来源的差异决定了公共数据主体要素公共性程度的差异。另一方面,公共数据的内容要素也存在公共性程度的差异,不仅包括公共安全、生态环境等纯公共产品,而且也包括教育医疗、公共交通等准公共产品。因纯公共产品供给活动产生的数据的公共性,显然大于因准公共产品供给活动而产生的数据。由此,基于公共数据内嵌主体要素与内容要素公共性差异的综合考量,可将其概括分为政务数据、公共非营利主体数据、公共营利主体数据和授权经营主体数据四种类型,也即公共数据的外延可以通过上述四个层次的数据类

型予以呈现。

(一)政务数据

第一层次的公共数据是政务数据,指职能履行受公共财政保障的非营利公共管理主体因提供公共产品而产生的数据。由于涵摄了公共性最高的全部要素,政务数据是公共性程度最高的数据类型。现实来看,政务数据基本等同于普通公众观念意义上的政府数据,不仅包括各级行政机关以及具有行政管理职能的事业单位的数据,而且也应包括人大、政协、司法和监察等其他国家机关制作或获取的数据。公共数据作为涅槃于“政府信息”的概念,将行政机关的数据纳入其范畴,并无疑义。然而,行政机关之外的其他国家机关制作或获取的数据是否应纳入公共数据范畴,则存在较大争论。当前,少部分地方立法并未将此类数据纳入公共数据范围,^③也有论者明确将其排除在公共数据范围之外。^④对此,需明确的是,虽然行政机关之外的其他国家机关的数据存量规模较小且多关涉国家秘密与安全,在现有数字技术环境下的公共利用价值确实较小,但这并非将其排除于公共数据范畴的正当理由。根据“主体+内容”的公共数据认定标准,此类数据之数源主体和数据内容的公共性与传统政务数据并无实质不同,两者本质上系同一层次的公共数据,将其排除于公共数据范畴有违公共数据之事物本质与法理逻辑。而且,随着行政机关之外的其他国家机关信息化进程的加速和数据处理技术的跃迁,此类数据资源的规模会逐渐庞大,公共利用的可能性亦会逐渐提升,将其排除于公共数据范畴的现实基础终将不复存在。

(二)公共非营利主体数据

第二层次的公共数据是公共非营利主体数据,指公共财政资金或社会捐赠资金支持的公共非营利主体提供社会服务而产生的数据。相较政务数据,公共非营利主体数据是非管理性主体提供准公共产品而产生的数据,公共性程度较弱且存在数据治理主体与数源主体分离的现象。虽然如此,但公共非营利主体仍是中国除国家机关之外公共性程度最高的公共主体,且数量众多,情况复杂,构成政府发挥公共服务功能的主要载体,管理模式行政化程度较

高。^⑤因此,大量公共非营利主体数据也是宝贵的公共数据资源,只有纳入公共数据范畴才能充分开发利用,实现数据价值的最大化。现实来看,此类数据主要是《民法典》规定的非营利性法人制作或获取的数据,纳入公共数据范畴基本无异议。申言之,公共非营利主体数据之数源主体包括如下几类:一是基金会,如中国红十字会、中国残疾人福利基金会等。二是公益服务类事业单位,如公立学校等教育事业单位、科研院所等科技事业单位、公立医院等卫生事业单位、报刊杂志社等文化事业单位、体育训练中心等体育事业单位等。三是社会团体,如中华全国总工会、中华全国妇女联合会、作家协会等。四是社会服务机构,如民办非营利性学校、民办非营利性医院、民办非营利性福利院等。

(三)公共营利主体数据

第三层次的公共数据是公共营利主体数据,指生产经营活动所需资金主要来源于公共财政资金的公共营利主体因提供准公共产品而产生的数据,其公共性体现为公共财政资金支持 and 提供准公共产品两个方面。相较前两类公共数据,公共营利主体数据不仅存在数据治理主体与数源主体相互分离的现象,而且产生数据的行为活动具有营利性目的,是公共性程度更弱的数据类型。现实来看,公共营利主体主要是提供电力、水务、燃气、通信、公共交通、城市基础设施建设等公共服务的营利法人性质的国有企业,比如国家电网、中国电信和各城市的自来水、燃气公司和公交集团等。由于数源主体的营利性法人性质,公共营利主体数据作为公共数据存在一定的争议,在较早的地方公共数据立法文本中,公共数据概念规范大多并未涵摄公共营利主体数据。其中,既有根本未涉及此类数据的立法文本,如《吉林省促进大数据发展应用条例》;也有通过附则中准用条款将其纳入公共数据治理范围的立法体例,比如《江西省公共数据管理办法》第3条和第45条之规定,再如《成都市公共数据管理应用规定》第2条和第37条之规定,亦如《上海市公共数据开放暂行办法》第3条和第47条之规定。不过,当前将公共营利主体数据直接纳入公共数据概念规范已成为主流立法趋势。如新近颁布的《广东省公共数据管理办法》第2条、《浙

江省公共数据条例》第3条、《上海市数据条例》第2条等。法理而言,公共营利主体数据作为由公共财政资金支持的主体提供准公共产品而产生的数据,符合“主体+内容”双重公共性的公共数据认定标准,理应纳入公共数据范畴。实践中,公共营利主体数据的认定尤需注意如下两方面,一是并非所有公共财政资金支持的国有企业制作或获取的数据都能够作为公共数据,而只有生产经营准公共产品的公共国有企业生产的数据才可能成为公共数据;二是公共营利主体制作或获取的数据也并非全部都能作为公共数据,只有因提供准公共产品而产生的数据才具有作为公共数据的合理性,反之,非因提供公共服务产生的数据需排除于公共数据范围。此乃公共数据之内容要素公共性的必然要求。

(四)授权经营主体数据

第四层次的公共数据是授权经营主体数据,指政府授权经营的非公共资金支持的营利性主体因提供准公共产品而产生的数据,即授权经营主体数据。相较其它类型的公共数据,授权经营主体数据同时存在治理主体与数源主体分离、产生数据活动性质的营利性以及主体要素的非直接公共性等特征,构成公共性最弱的公共数据类型。现实中,授权经营主体数据主要对应着民办营利性学校、民办营利性医院以及社会资本主导的基础设施与公用事业特许经营主体等提供准公共产品且市场准入需经政府特别授权的营利性主体产生的数据。因主体要素公共性的缺失,此类数据是否应当纳入公共数据范畴存在较大争议。就地方公共数据立法来看,只有《深圳经济特区数据条例》《济南市公共数据管理办法》《安徽省大数据发展条例》等少部分立法文本明确将政府授权经营主体的数据纳入公共数据范畴。表面来看,政府授权经营主体提供公共服务以营利为目的,其产生的数据作为公共数据仅因其内容要素涉及准公共产品。然而,事实并非如此。正如前文所述,授权经营主体本质上是替代政府向社会提供公共产品,需受到政府的严格干预和限制,并在一定程度上承担着社会治理责任,在享受该角色带来的经济利益的同时也需负担一定程度的数据处理义务。换言之,准公共产品的供给实际上并未完全市

场化,提供准公共产品的虽然是非公共性的市场主体,但市场准入和经营需要受到政府严格的审查批准和监管限制。在此意义上,虽然此类主体不具有直接公共性,但却因政府的特别授权而具有一定程度的间接公共性,从而在形式逻辑上仍然符合“主体+内容”双重公共性的公共数据认定标准。至于负担数据处理义务与其商业秘密等法益的冲突,可通过高阶法益优先保护的规则予以协调,并不构成阻却授权经营主体数据作为公共数据的正当理由。同时需强调的是,由于准公共产品界限的模糊性,授权经营主体数据的认定较为依赖政府授权,即政府主管部门的审查批准或特许经营。亦即,只有获得政府审查批准或特许经营的主体制作或获取的数据方才属于公共数据范畴。

余论:公共数据治理的类型化路径

公共数据的概念范畴是公共数据治理这一鲜明时代议题的基石。作为数字经济时代由“政府信息”涅槃新生的概念,立法实践与理论学说对其内涵外延产生了不同的认知,并造成公共数据治理实践的诸多问题。在梳理立法文本与理论学说的基础上,本文批判性地反思了当前“归属标准”和“用途标准”的缺陷,并基于公共数据之事物本质提出“主体+内容”双重公共性的公共数据认定标准,即只有数源主体和数据内容本身都表征公共利益的数据才具有作为公共数据的法理正当性。据此,公共数据应当是职能履行受公共财政资金或社会捐赠资金支持公共性主体及政府授权生产经营准公共产品的市场主体,因履行公共管理职责或提供社会公共服务而制作或获取的数据。根据公共性程度的差异,可将“主体+内容”双重公共性标准框定的公共数据范畴分为公共性程度渐次降低的政务数据、公共非营利主体数据、公共营利主体数据和授权经营主体数据四类。至此,公共数据的概念范畴已经清晰可见。

然而,需注意的是,本文的论述事实上也附带出一个新的学术议题。既然公共数据的主体要素与内容要素表征的公共性程度存在差异,即公共数据作为规模庞大的数据资源池存在异质性和层次性,那么公共数据治理就无法适用整齐划一的治理规则。现实来看,部分类型的公共数据未纳入特定地方立

法文本的概念规范,也正是源于立法者对其负担某项公共数据处理义务能力的担忧和质疑。理性而论,公共数据负担共享、开放、安全管理等义务是其本身作为公共利益表征的合理逻辑演绎,而公共数据内嵌主体要素与内容要素公共性程度差异则表明不同类型的公共数据负担上述义务的能力并不相同。基于公共数据的异质性和层次性,公共数据治理应当遵循类型化的路径,为各类公共数据设置不同的治理规则,使其负担与其公共性特征相妥适的义务和责任。概括来看,政务数据作为公共性最高的数据类型,应当负担完整的数据共享、开放和管理义务。公共非营利主体数据因存在治理主体与数源主体分离的现象,数源主体的数据管理能力存在一定程度的瑕疵,应当负担数据共享、开放义务和较高度度的数据管理义务。公共营利主体数据同时存在治理主体与数源主体的分离和数源主体的营利性特征,可以只负担数据共享义务和适度的数据管理义务。授权经营主体数据作为公共性最弱的公共数据,不仅存在治理主体与数源主体分离现象和产生数据之行为性质的营利性,而且数源主体的直接公共性亦不复存在,只需负担数据共享义务。^⑧当然,公共数据类型化治理是一个复杂的研究问题,非此寥寥数言可以阐释,尚有待进一步的学术努力!

注释:

①参见戴昕:《数据界权的关系进路》,载《中外法学》2021年第6期,第1577页。

②参见齐英程:《作为公物的公共数据资源之使用规则构建》,载《行政法学研究》2021年第5期,第143页。

③余筱兰:《公共数据开放中的利益冲突及其协调——基于罗尔斯正义论的权利配置》,载《安徽师范大学学报》(人文社会科学版)2021年第3期,第84页。

④郑大庆、黄丽华、郭梦珂、卢向华:《公共数据资源治理体系的演化模型:基于整体性治理的建构》,载《电子政务》2022年第5期,第80页。

⑤袁康、刘汉广:《公共数据治理中的政府角色与行为边

界》,载《江汉论坛》2020年第5期,第121页。

⑥参见齐英程:《作为公物的公共数据资源之使用规则构建》,载《行政法学研究》2021年第5期,第143-144页。

⑦参见郑春燕、唐俊麒:《论公共数据的规范含义》,载《法治研究》2021年第6期,第74页。

⑧参见齐英程:《作为公物的公共数据资源之使用规则构建》,载《行政法学研究》2021年第5期,第143页。

⑨[德]托马斯·M.J.默勒斯:《法学方法论》(第4版),杜志浩译,北京大学出版社2022年版,第484页。

⑩参见[英]迈克·费恩塔克:《规制中的公共利益》,戴昕译,中国人民大学出版社2014年版,第1页。

⑪参见张千帆:《“公共利益”的构成——对行政法的目标以及“平衡”的意义之探讨》,载《比较法研究》2005年第5期,第8页。

⑫参见胡凌:《论地方立法中公共数据开放的法律性质》,载《地方立法研究》2019年第3期,第18页。

⑬参见江帆:《经济法的价值理念和基本原则》,载《现代法学》2005年第5期,第120-121页。

⑭比如,1875年科赫诉联邦政府案(Kohl v. United States)、1896年密苏里太平洋铁路公司诉内布拉斯加州案(Missouri Pac. Ry. Co. v. State of Nebraska)等。

⑮Thomas M. Cooley, A Treatise on the Constitutional Limitations Which Rest upon the Legislative Power of the States of the American Union, Boston: Little Brown Company, 1868, pp.585-586.

⑯参见楼利明:《论“公共利益”与“公共使用”》,载《武汉大学学报》(哲学社会科学版)2006年第5期,第614-615页。

⑰参见袁康、刘汉广:《公共数据治理中的政府角色与行为边界》,载《江汉论坛》2020年第5期,第121页。

⑱参见胡凌:《论地方立法中公共数据开放的法律性质》,载《地方立法研究》2019年第3期,第6页。

⑲参见郑春燕、唐俊麒:《论公共数据的规范含义》,载《法治研究》2021年第6期,第76页。

⑳参见刘兴华:《德国“债务刹车”的缘起、机制与功效》,载《外国问题研究》2021年第1期,第61页。

㉑参见刘尚希、傅志华、李成威、陈龙、申学锋:《全面认识财政是国家治理的基础和重要支柱——学习习近平总书记关于财政问题的重要论述》,载《财贸经济》2022年第2期,第14页。

㉒《中华人民共和国审计法》第24条第2款规定:“审计机

关对政府部门管理的和其他单位受政府委托管理的社会保险基金、全国社会保障基金、社会捐赠资金以及其他公共资金的财务收支,进行审计监督。”《中华人民共和国公益事业捐赠法》第20条第1款规定:“受赠人每年度应当向政府有关部门报告受赠财产的使用、管理情况,接受监督。必要时,政府有关部门可以对其财务进行审计。”

②③参见徐德骥、洪婧茹:《谨慎对待城市“公私合营”与公共数据私人控制——评〈数字化世界的智慧城市〉》,载《国际新闻界》2020年第10期,第160页。

②④参见秦颖:《论公共产品的本质——兼论公共产品理论的局限性》,载《经济学家》2006年第3期,第79页。

②⑤参见寇铁军、钟永圣:《经济学理性主义与财政学基础理论的发展》,载《财政研究》2008年第7期,第3-4页。

②⑥参见于君博、林丽:《产品性质与公共产品供给绩效——

应急性的凸显及其影响》,载《经济社会体制比较》2021年第3期,第93页。

②⑦参见许可:《数据权利:范式统合与规范分殊》,载《政法论坛》2021年第4期,第93页。

②⑧比如,《广东省公共数据管理办法》《无锡市公共数据管理办法》等。

②⑨参见赵家兵:《公共数据归属政府的合理性及法律意义》,载《河南财经政法大学学报》2021年第1期,第13页。

③⑩参见李文钊、董克用:《中国事业单位改革:理念与政策建议》,载《中国人民大学学报》2010年第5期,第139页。

③⑪关于公共数据治理的类型化路径。作者在另文中有相对完整的讨论。具体参见沈斌、黎江虹:《论公共数据的类型化规制及其立法落实》,载《武汉大学学报》(哲学社会科学版)2023年第1期,第67-77页。

Identification Criteria and Type System of Public Data

Shen Bin

Abstract: As the cornerstone of public data governance, the concept category of public data has not reached the legislative unity and academic consensus. Whether the attribution criteria with the publicity of data subject elements as the core adopted by the current legislation, or the use criteria with the public use value of data content as the core proposed by the academic community, all have defects. Based on the essence and value characteristics of public data, the definition of the category of public data should adhere to the dual publicity criteria of "subject+content", that is, only data that both the data source subject and the data content itself represent the public interest can be regarded as public data and bear corresponding obligations of data processing. The publicity of the subject elements is characterized by the guarantee of public funds or the operation authorized by the government, and the publicity of the content elements is possessed as a by-product of the supply of public goods. According to the difference between the publicity of subject elements and content elements, public data can be divided into four categories: government data, public non-profit subject data, public profit subject data and authorized operation subject data.

Key words: Public Data; Public Interest; Public Goods; Identification Criteria; Dual Publicity