

公共数据的产权运行机制与技术方

郑宇^{1,2,3}

1. 京东城市(北京)数字科技有限公司, 北京 100176;
2. 京东智能城市研究院, 北京 100176;
3. 西南交通大学计算机与人工智能学院, 四川 成都 611756

摘要

公共数据的要素化既有迫切需求和巨大价值, 又具备良好基础和实施保障。为了实现公共数据的要素化, 数据要素智能构建技术与公共数据产权运行机制需要协同创新和融合发展。梳理了公共数据的分类和参与主体, 提出一套公共数据产权运行机制, 包括公共数据产权结构性分置制度、公共数据确权授权机制和个人信息授权机制, 并搭建一套数字化系统来承载该机制, 以建立健全数据要素各参与方的合法权益, 促进公共数据在政府各部门、各层级、各主体之间的安全、合规、有序流通。

关键词

数据要素; 公共数据; 数据产权运行机制; 数据产权结构性分置; 公共数据确权授权

中图分类号: TP399

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2024055

Mechanisms and techniques for operating public data right

ZHENG Yu^{1,2,3}

1. JD Intelligent Cities Technology Co., Ltd., Beijing 100176, China
2. JD Intelligent Cities Research, BDA, Beijing 100176, China
3. School of Computing and Artificial Intelligent, Southwest Jiaotong University, Chengdu 611756, China

Abstract

Data as a factor of production (i.e. data factors) is the key momentum for boosting digital economy. Turning public data into data factors needs the synergy between technology and mechanisms. In this paper, we categorize public data into two types and define the four main parties participating in the generation of public data factors. Dealing with these two types of data and four parties, we propose a data-right operational system, which is comprised of a data-right allocation regulation and a set of data-right authorization mechanisms. While ensuring the security of the public data and the compliance with data-right regulations, the system can protect the rights of each parties, enabling them to use data more effectively and efficiently, and thus unlock the power of public data.

Key words

data factors, public data, data-right operational mechanism, data-right allocation regulation, data-right authorization mechanism

0 引言

数据要素对于推动经济增长、优化资源配置、提升产业链现代化水平具有重要的战略意义。2022年12月,《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》(简称“数据二十条”)正式发布,以数据产权、流通交易、收益分配、安全治理为重点。随后国家数据局等17个部门联合印发《“数据要素×”三年行动计划(2024—2026年)》,提出“更好发挥政府作用,扩大公共数据资源供给”,并细化了数据要素在城市治理、应急管理、文化旅游等12个领域的工作重心。至此,攻关数据要素的大幕正式拉开。

按照数据相关权益归属的不同,“数据二十条”将数据分为公共数据、企业数据、个人数据,并强调对公共数据“加强汇聚共享和开放开发,强化统筹授权使用和管理,推进互联互通,打破‘数据孤岛’”。公共数据是各级党政机关、企事业单位依法履职或提供公共服务过程中产生的数据,此类数据覆盖面宽,应用场景丰富^[1]。多年的政务信息化工作为公共数据形成了较好的治理基础,且政府对于该类数据的管理与开放具有较大的推动力和掌控力。因此,将公共数据作为数据要素工作的重心既有迫切需求和巨大价值,也具备更好的工作基础和实施保障^[2]。

要实现构建公共数据要素的战略目标,需要技术和机制的协同创新。一方面,要基于智能技术建立可共享互认的公共数据资源体系,要使数据和应用分离,要能自动产生、更新和连接数据要素,从而形成规模效应,发掘数据价值^[3-5]。另一方面,需要建立保障权益、合规使用的数据产权制度和安全可控、弹性包容的数据要素治

理制度,这样公共数据要素才能合规、高效地流通使用^[6]。这些制度要通过技术落实到具体工作环节和应用场景中,而不是仅仅停留在文字层面。

要实现公共数据的合规高效流通,首先要解决其在政府内部不同部门、不同层级之间的共享和流通问题,之后才有可能实现更大范围(如商业市场)的循环。公共数据的产生、流通、使用和治理环节涉及居民、政务基层工作者、数据管理部门(如大数据局或政数局)以及政府其他业务部门(如卫健委、应急局、市场监管局和街道)等多个主体。当前,这些主体在数据产权分置、公共数据确权授权、个人信息数据确权授权及数据要素参与方合法权益保护方面存在以下挑战。

(1) 基层工作者拥有的数据权限不明确,安全和效率的平衡点有待探寻

基层工作者既是数据的使用者,也是数据的创造者,他们在工作过程中产生了大量数据,但由于数据权属不清、授权机制缺失,在一个业务场景中积累的数据无法在其他业务中使用^[7]。同时,基层工作者经常为居民服务,会接触到不少居民个人信息,如果处理不当,容易引发数据安全和合规问题,工作者个人不能保留这些数据。因此,基层工作者一直在贡献数据却无数可用的尴尬现象普遍存在,基层工作者的工作效率大打折扣^[8]。产数和用数的矛盾如何化解?安全和效率的平衡点如何找到?为保障基层工作者作为数据要素参与方的合法权益,急需建立健全数据产权制度和安全治理机制。

(2) 业务部门和数据管理部门之间的数据管理责权不清晰,部门之间的管理边界有待细化

政府委办局在开展业务的过程中已经搭建了不少信息系统,并产生了大量的公共数据。有些业务拥有独立的系统和数据

库,需要将数据进一步汇聚到一个中心化数据平台,之后才能进一步在不同部门之间共享和流通。还有部分城市已经建设了统一的数据平台,并在这些平台上开展业务,数据也存放在一起。一般情况,大数据局作为数据能力的支撑部门,负责统一数据平台的运行和维护。但大部分数据由业务部门在开展业务时产生和使用,这些数据在产生、供给、流通和使用环节的权责如何划分还存在很多不清晰的地方,数据授权机制不够精准高效,数据治理和供给存在瓶颈。例如,业务部门拥有自身业务数据的何种产权?大数据局对业务数据有何种权限,对加工后的数据又拥有何种权限?不同业务部门之间如何使用彼此的数据?如果部门之间都进行点对点的相互授权,那大数据局在其中又扮演何种角色?这样的授权机制的效率能满足业务开展的实际需求吗?反之,如果所有的数据管理职责都交给大数据局,或者各部门都向大数据局提出数据治理的需求,大数据局将承受繁重的工作,很有可能成为数据流通的瓶颈^[9]。

(3)居民个人信息数据确权 and 授权模糊,急需构建授权和收益的桥梁

公共业务需要基于个人信息为居民提供精准、高效的公共服务,合理的授权能减少不必要的证明材料的提供和审核处置流程,避免居民在不同政务服务中重复填报信息。但当前精细化、人性化的授权机制仍然缺失^[10]。一方面,统一、过度授权的现象引起居民对个人信息安全的担忧;另一方面,存在过度谨慎、完全放弃的情况,各部门对授权问题

避而远之,导致居民无法享受数据的红利。如何让居民既能牢牢掌握个人数据的权属,又能通过精准合理授权享受数据红利,这是当前急需解决的问题之一。

为解决以上问题,本文提出一套公共数据产权运行机制,包括公共数据产权结构性分置机制、公共数据确权授权机制和个人信息授权机制,并搭建一套数字化系统来承载该机制,通过技术与机制的协同创新来建立健全数据要素各参与方的合法权益,从而促进公共数据在政府各部门、各层级、各主体之间的流通。

1 公共数据的种类

如表1所示,按照形成的方式,可将公共数据分为公共业务数据和公共数据要素两大类。进一步,根据是否涉及个人信息,可将公共业务数据分为业务管理数据和个人信息数据,将公共数据要素分为普通数据要素和个人信息要素,共4类数据。

(1) 公共业务数据

公共业务数据是各级党政机关、企事业单位依法履职和提供公共服务过程中产生的业务数据。公共业务数据包括两个子类。一类是业务管理数据,如工作执行过程、服务审批记录、事件处置结果、居民反馈建议、信息统计报表等。这类数据不涉及居民的个人信息,产生方式包含政务系统自动生成和政务人员输入创建。另一类是居民在使用公共服务过程中提供的个人信息,

表1 公共数据的分类

数据种类	不涉及个人信息		涉及个人信息	
	细分类别	产生方式	细分类别	产生方式
公共业务数据	业务管理数据	政务系统自动生成; 政务人员输入创建	个人信息数据	居民自身填报; 委托他人填写
公共数据要素	普通数据要素	智能系统自动形成; 人机协同治理产生	个人信息要素	系统感知提示; 居民逐个授权

如证件号码、出生年月、联系方式、家庭住址、工作单位、健康状况、家庭成员及个人物产等信息,产生方式包括居民自身填报和委托他人(包括政务人员)填写。业务数据中的个人信息只能服务于产生该数据的业务系统,技术上不能、制度上无法在不同应用间自动共享。注意,个人信息数据并不是“数据二十条”中划分的“公共数据、企业数据和个人数据”中的个人数据,以上3类数据均有可能承载个人信息。

(2) 公共数据要素

公共数据要素是大数据局等数据管理部门对公共业务数据进行加工生产后形成的可在不同部门间共享流通的数据。公共数据要素包括两个子类。一类是基于业务管理数据构建的普通数据要素,如组织、物件、地点和事件的常用属性和关联关系,用于解决不同政务部门间信息共享问题。该类要素通过大数据局管理的智能信息平台自动产生、更新和相互连接^[5],或针对已经形成的历史业务数据,利用数据治理工具以人机协作的半自动化方式转化产生。另一类

是基于各种个人信息数据形成的个人信息要素,用于解决居民在不同公共服务中反复填报信息和提供冗余证明的问题。个人信息数据属于某个具体的业务,而个人信息要素能在不同应用间共享,可支撑不同的业务。居民在使用公共服务系统时,通过系统自动检测提示、居民确认授权的方式将个人信息逐一加入个人信息要素中。

2 公共数据产权运行机制

2.1 总体框架

公共数据产权运行机制的总体框架如图1所示,其由法理依据、管理机制和技术保障三部分构成。

(1) 法理依据

法理依据包含“数据二十条”、《数字中国建设整体布局规划》和《“数据要素×”三年行动计划(2024—2026年)》等国家出台的重要战略规划,为制度管理和

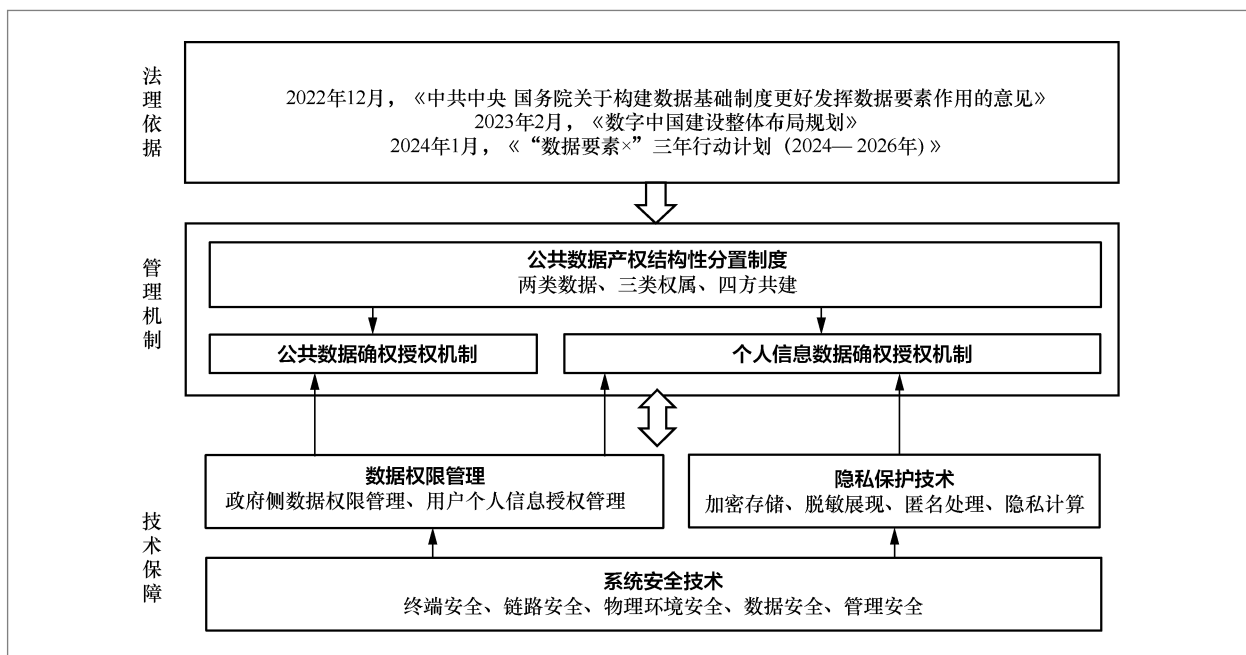


图1 公共数据产权运行机制

技术保障的设定提供理论基础和指导方针。其中,“数据二十条”以“建立保障权益、合规使用的数据产权制度”为目标,在第三、四、六和七条分别强调“探索数据产权结构性分置制度、推进实施公共数据确权授权机制、建立健全个人信息数据确权授权机制、建立健全数据要素各参与方合法权益保护制度”。《数字中国建设整体布局规划》以畅通数据资源大循环为目标,从“构建国家数据管理体制机制,健全各级数据统筹管理机构”和“释放商业数据价值潜能,加快建立数据产权制度”两个维度部署工作。《“数据要素×”三年行动计划(2024—2026年)》第二十条中提出“研究数据资源持有权、数据加工使用权、数据产品经营权等分置的落地举措”。

(2) 管理机制

管理机制包括公共数据产权结构性分置制度、公共数据确权授权机制和个人信息数据确权授权机制,确保数据安全、合规、有序流动,保障数据要素各参与方的合法权益。其中,公共数据产权结构性分置制度对数据类型、权利内容和权利主

体进行结构性划分。公共数据确权授权机制明确政府不同主体对不同数据类型的权利内容,以及不同主体之间的数据授权机制(如业务部门对大数据局和基层工作者等主体的数据授权)。个人信息数据确权授权机制明确个人对其在公共服务中提供的信息所拥有的权属内容,以及个人对不同政府主体的数据授权方式。

(3) 技术保障

技术保障包含数据权限管理、隐私保护技术和系统安全技术。其中,数据权限管理承载公共数据确权和授权机制,并与隐私保护技术共同支撑个人信息数据确权授权机制,保障各方合法权益,确保数据合规、有序流通。系统安全技术确保数据在产生、加工、流通、使用等过程中的信息安全,为数据的安全治理和合规使用提供坚实的基础保障。

2.2 公共数据产权结构性分置制度

如图2所示,从数据类型、数据权属和参与主体3个维度构建了公共数据产权立

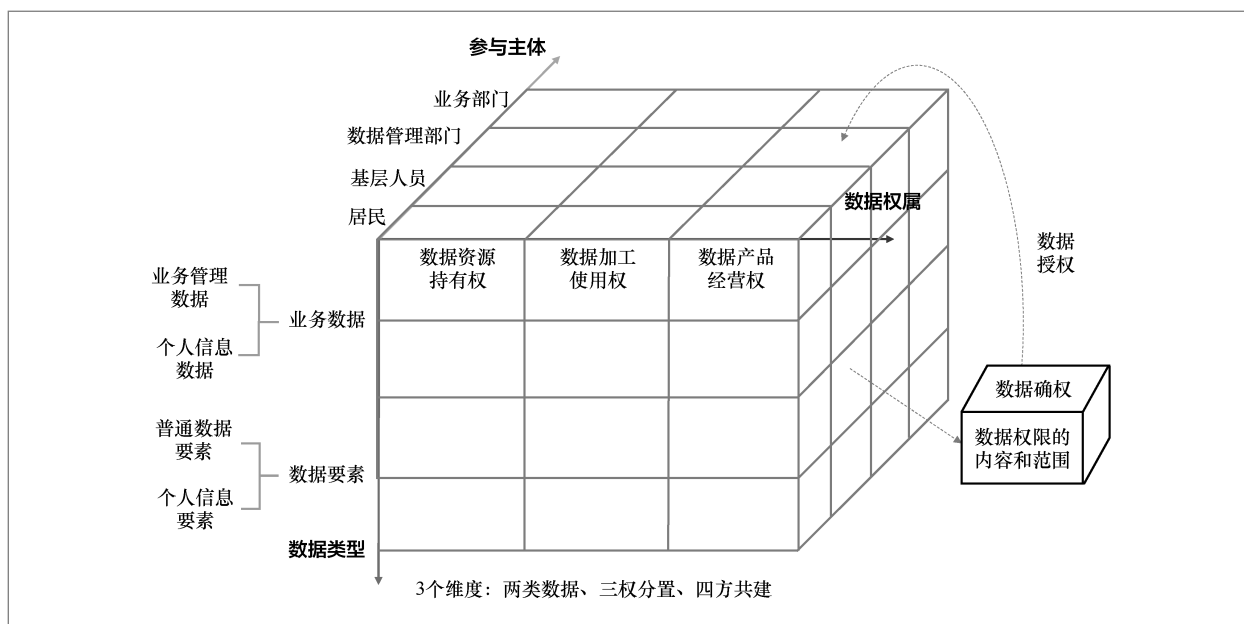


图2 公共数据产权分置制度方案

方体。结构化的产权分置制度使不同主体对不同数据拥有不同的数据权属,为后续精细化的数据确权授权机制打下基础。

(1) 公共数据类型

公共数据可以分为公共业务数据和公共数据要素两大类,并进一步细分为四小类(详见第1节)。

(2) 数据权属

根据《“数据要素×”三年行动计划(2024—2026年)》,数据权属分为数据资源持有权、数据加工使用权和数据产品经营权。数据资源持有权通常涉及数据的获取、存储、使用、更新、共享和保护等方面的决策和操作,即在公共数据场景中对数据能改、能删、能授权。数据加工使用权是对数据收集、存储、使用、加工、传输、提供的权利,数据加工包括抽取、清洗、分析、统计、转换、运算和进一步挖掘,但不能修改、删除数据,也不能进行数据授权。数据产品经营权是指相关机构对加工形成的数据产品依法占有、使用、收益和依法支配的权利。

(3) 参与主体——四方共建

公共数据的参与主体包括居民、基层工作者、数据管理部门和业务部门。针对不同数据,这些主体可具有以下4种身份中的一种或多种:数据所有者、数据生产者、数据使用者和数据管理者。例如:居民在使用公共服务的过程中既是个人数据的所

有者,也是该数据的生产者、使用者和管理者;基层工作者是业务管理数据的生产者和使用者,也可能是个人信息的产生者(如政务人员代居民填报)和管理者;业务部门是业务管理数据的所有者、产生者、使用者和管理者,也是公共数据要素的使用者;数据管理部门是业务管理数据和个人信息数据的使用者和管理者。4类主体协同完成公共数据的生产、加工、存储、更新、使用和流通,实现公共数据要素的四方共建。

如图2所示,针对两种公共数据类型,构建了三权分置和四方共建的公共数据产权制度,在每个小立方体中可以明确具体权限的范围和内容,实现数据确权。进一步,一个主体也可以根据数据确权内容,合理合规地将某些权限授予其他主体(如居民将个人信息数据的加工使用权授予业务部门和数据管理部门),以实现效率与安全的平衡,理清数据管理边界,并搭建授权和收益的桥梁。

2.3 公共数据确权授权机制

(1) 公共数据确权机制

表2展现了不同主体对不同种类的公共数据拥有的数据权利,部分权利需要经过合理的授权。此外,数据权利的大小会因业务部门的层级、管辖范围和业务场

表2 公共数据的确权

数据类型		数据管理部门	业务部门	基层工作者	居民
业务数据	业务管理数据	数据加工使用权	数据资源持有权; 数据加工使用权	数据加工使用权	—
	个人信息数据	数据加工使用权	数据加工使用权	数据加工使用权	数据资源持有权
数据要素	普通数据要素	数据加工使用权; 数据产品经营权	(基于本部门数据产生的要素) 数据资源持有权; 数据加工使用权	—	—
	个人信息要素	数据加工使用权	(其他部门数据产生的要素) 数据加工使用权; 数据加工使用权	—	数据资源持有权

景而不同。该方案综合考虑了国家法律法规、战略规划文件和政府运行机制,能应对引言部分提及的三大挑战。

- **业务管理数据的确权:**业务部门拥有本部门开展业务时产生的业务管理数据的数据资源所有权和数据加工使用权,能对该数据进行修改、删除和进一步授权等操作,也能在本部门开展的其他业务中使用该数据。经业务部门授权后,数据管理部门和基层工作者拥有该数据的加工使用权。基层工作者利用智能系统帮助业务部门在中心数据平台直接产生业务数据,而不是先在本地产生数据再通过系统提交上传至中心数据平台。因此,基层工作者并不拥有数据资源所有权,但可以从中心数据平台调取、使用这些业务数据来开展工作。该模式同样适用于个人信息数据。

- **个人信息数据的确权:**居民对其在某个公共服务中提供的个人信息数据拥有数据资源所有权,经用户授权后,在该服务中,业务部门、数据管理部门和基层工作者拥有对用户个人信息数据的加工使用权;离开该服务,这些主体不再拥有对用户个人信息数据的加工使用权。基层工作者帮助产生和使用个人信息数据的方法跟业务管理数据的模式一致,即在后台直接产生,在前端调用。

- **普通数据要素的确权:**数据管理部门对普通数据要素拥有加工使用权和数据产品经营权。业务部门对数据管理部门基于本部门业务管理数据加工产生的普通要素数据拥有数据资源所有权和数据加工使用权;对基于其他业务部门数据产生的普通数据要素仅拥有数据加工使用权,且该权利的具体范围和大小受数据管理部门的管控。基层工作者不直接拥有普通数据要素的加工使用权。业务部门需要基于业务需求,利用其对普通数据要素的数据加工使用权,并根据该部门的权限范围使用普

通数据要素来补全业务数据中的不足。之后,该业务部门再将补全后的业务数据及数据的加工使用权赋予基层工作者,基层工作者在此业务中的具体权限范围和大小受该业务部门的管控。

- **个人信息要素的确权:**居民拥有该数据的数据资源所有权,并通过逐一确认的方式将自己的个人信息要素的加工使用权授予政府侧的数据管理部门和业务部门。后续数据管理部门根据各部门的管辖权限和业务需求对个人信息要素进行精细化权限管控,以满足不同层级部门和工作人员的用数需求。一旦用户删除了某项个人信息要素,数据管理部门维护的个人信息要素也会被自动删除,数据资源所有权永远掌握在居民手中。

(2) 公共数据的授权机制

图3展现了业务管理数据和普通数据要素的授权机制。业务部门拥有本部门开展业务产生的业务管理数据的数据资源所有权,通过统一授权的方式(即不用每个部门针对每项业务分别授权),将这些业务管理数据的加工使用权赋予政府的基层工作者、数据管理部门和其他业务部门。进一步,业务部门通过“业务权限管控”设置基层工作人员对业务管理数据的权限大小,数据管理部门利用“部门权限管控”设置业务部门对非自身贡献的普通数据要素的权限大小。

- **业务权限管控:**如**图3**流程①所示,业务部门根据业务需求,基于信息最小泄露原则,管控基层工作者对业务管理数据的数据加工使用权的范围和大小。例如,部分服务只允许基层工作者查看自己审批的事项,部分服务可以查看同社区所有工作者的审批记录;部分业务可以共享转让,而部分业务只能自己独立完成。因此,业务管理数据的统一授权,仅代表数据权属种类的一致性,但权限的范围和大小会

因具体人员和场景的不同而不同。

● 部门权限管控：数据管理部门默认拥有所有业务管理数据的加工使用权，对不同部门产生的业务管理数据进行集中加工，形成可在不同应用间共享的普通数据要素（如组织、物件、地点、事件的基本信息和相互关系）。如图3流程②所示，数据管理部门根据用数部门（如业务部门2）的组织层级、部门职责和业务性质来精准控制其对普通数据要素的权限大小。如图3流程③所示，业务部门2在获得的权限范围内，又根据业务权限管控进一步限定在本业务中基层工作者的权限范围和大小。

不同业务部门之间无须进行点对点的数据授权，一方面该授权方式效率低下，沟通成本高；另一方面，一个部门的业务数据无法直接被另一个部门使用，需要进行加工处理。加工处理工作应由数据管理部门集中、自动完成，而不应让每个业务部门分别手动处理数据。

2.4 个人信息数据和要素确权授权机制

第2.3节详细介绍了个人信息数据和个人信息要素的确权，本节重点介绍个人对这两类数据的授权机制。图4左边部分描述了个人信息数据在单个服务中的授权方式，右边部分展现了个人信息要素在跨多个服务中使用的授权方式。

(1) 个人信息数据的授权

如图4流程①所示，个人（如居民）在使用某项公共服务时，对本人在该服务中提供的个人信息拥有数据资源持有权。个人可将此服务中的个人信息数据的加工使用权统一授予政府（包括与该服务相关的业务部门、数据管理部门和基层人员），但这个授权仅限于该服务内部，这些个人信息数据不能直接被其他公共服务使用。

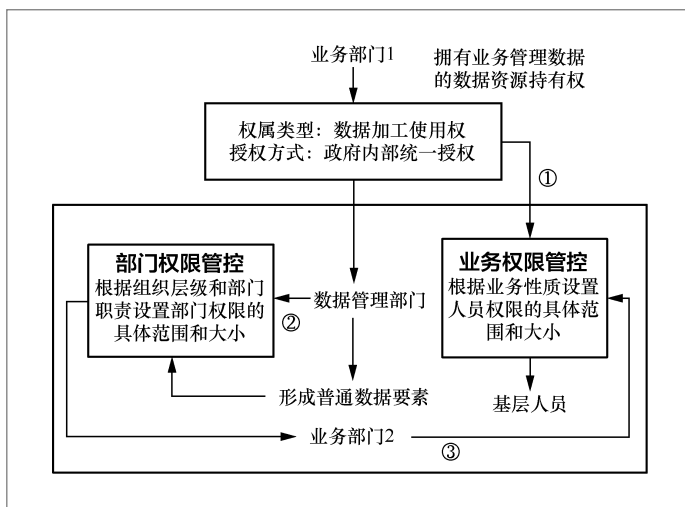


图3 业务管理数据和普通数据要素的授权

此处的“统一”是指对单个服务中的多项个人信息字段（如电话、住址等）的一同授权（与个人信息要素的“逐一”授权不同），以及对政府内部多个主体的一致授权（即在一个服务的用户授权提示中集中表达）。如图4流程②所示，业务部门获取用户个人信息数据在该服务中的加工使用权后，根据业务的需求和信息最小泄露原则，通过“业务权限管控”进一步设置基层人员的权限大小。因此，基层工作者对一个服务中个人信息数据的加工使用权由用户个人授予，但权限的具体范围和大小由业务部门管控。

(2) 个人信息要素的授权

当政府希望利用居民提报过的个人信息来优化服务流程、减少重复填报时，需要将个人信息数据加工为个人信息要素，这样个人信息才能在不同应用间共享。此时的授权方式既要严格、精准，但也需考虑用户体验，减少用户的授权负担。同时，个人信息也在不断变化，个人信息要素的更新无论从数据资源持有权的法理依据来看，还是从现实工作量及可行性来看，只能让用户自己操作，政务工作者和数据管理部门皆不能代

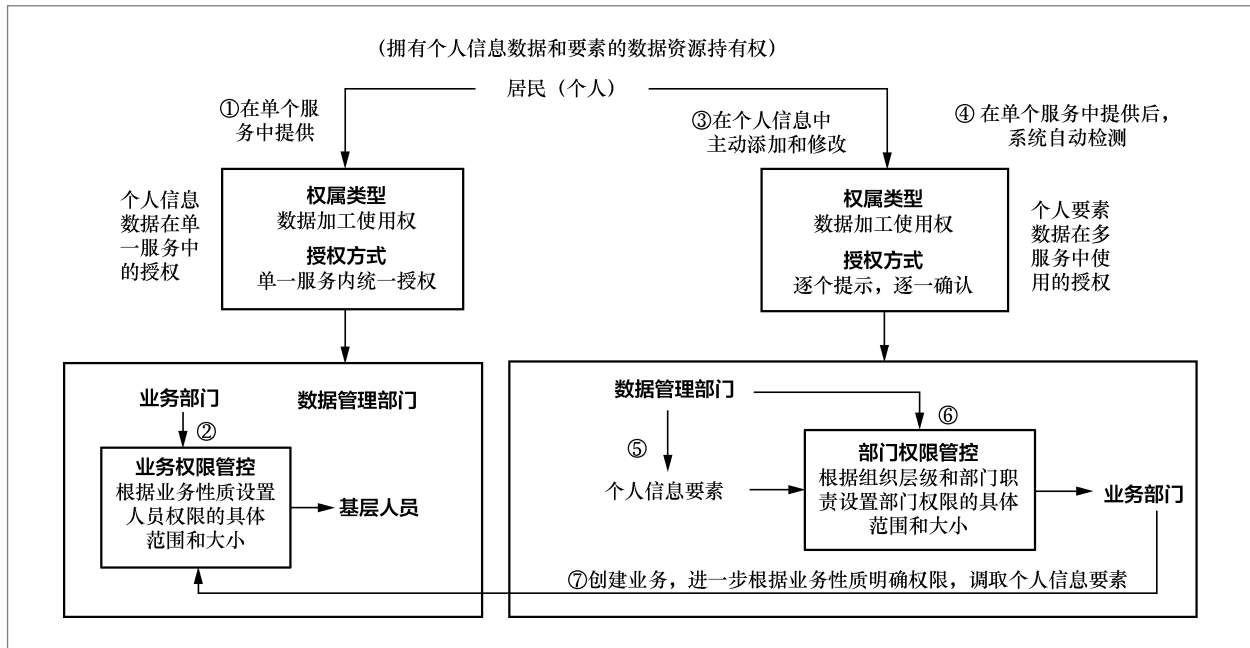


图4 个人信息数据和要素的授权机制

劳。因此，对个人信息要素的授权是一种持续动态的过程，需要系统和技术的支撑以及居民和数据管理部门的协作。

如图4右部分所示，个人可通过以下两种方式对个人信息要素授权。

第一种，如图4流程③所示，在公共服务系统的用户端中，每个用户都有一个管理和维护个人信息的界面（例如，App中通常叫“我的”）。个人信息包括姓名、证件号、电话、家庭住址、家庭成员等。用户在此处主动对这些信息进行添加和修改时，系统会逐一提示授权，明确告知用户“为了给您提供更加优质、便捷的服务，减少您的重复填报，授权本系统在其提供的服务中使用该信息”。

第二种，如图4流程④所示，当个人在使用某项公共服务时，填报了新的个人信息（如新的家庭住址），系统会自动精准提示用户“您提供了新的家庭住址，是否要添加到您的个人信息中，以便在其他服务中再次使用”。如果用户选择同意，则该数据在前端界面中会被加入“个人信息”里，同

时在后端的中心化数据平台中形成个人信息要素，继而被其他公共服务使用。

在以上两种情况中，用户将个人信息要素的加工使用权精准、逐一授予政府，数据管理部门（按照职责认定）作为政府代表对该要素数据进行加工和管理，在后端的数据平台上形成与业务分离的个人信息要素（如图4流程⑤所示）。当其他业务部门在开展工作中需要调用个人信息要素时，基于信息最小泄露原则，如图4流程⑥所示，数据管理部门通过“部门权限管控”，按照需求部门的组织层级、部门职责和业务场景，设置其拥有的数据加工使用权限的范围和大小。

例如，根据组织层级，区级部门在其提供的公共服务中只能调用本区居民的个人信息要素；街道管委会只能调取本街道居民的个人信息要素。根据部门职责，卫健委可使用与个人健康状况相关的信息要素；民政局可使用与个人婚姻状况相关的信息要素。根据业务场景，与老人关怀相关的服务可调取家庭成员等个人信息要素。

一个部门的权限受以上3个维度的约束,即在用户授权政府使用个人信息要素的加工使用权后,区民政局在与老人关怀相关的公共服务中可调取本区居民的家庭成员信息。

如图4流程⑦所示,根据以上权限约束,业务部门的应用将所需的个人信息调入业务中,形成业务数据,并进一步按照业务权限管控的方法授权。基层工作者只有通过业务部门的具体业务才能间接使用个人信息,避免了直接管理和使用个人信息要素。因此,基层工作者拥有个人信息数据的加工使用权,没有也不需要个人信息要素的加工使用权。

2.5 技术保障

技术保障由数据权限管理、隐私保护技术和系统安全技术3个部分组成。由于探讨隐私保护和系统安全技术的文章较多,本文重点介绍公共数据权限管理,包括政府侧数据权限管理和用户个人授权管

理两部分。该技术可以解决公共数据跨管辖域和跨信任域流通的问题^[11]。

(1) 政府侧数据权限管理

如图5所示,政府侧数据权限管理由部门权限管控、业务权限管控和数据访问控制3个模块组成。

● 部门权限管控模块:如图5流程①所示,数据管理部门通过政府的组织通讯录获取业务部门的组织层级(如市-区-街-居),根据所在层级节点明确其管辖范围。例如,北京市卫健委属市级委办局,管辖全市居民的公共卫生业务;朝阳区民政局属于区级委办局,负责朝阳区居民的民政工作。如图5流程②所示,数据管理部门通过组织和业务共享的标签体系,如环境卫生、老人关怀、婚姻登记、房产管理、殡仪服务、幼儿教育等,明确组织职责及某个具体业务的性质。一个部门的职责用若干个组织标签来表示,部门职责源于政府侧的“部门三定”(依照严格的程序制定和审核),数据管理部门只是执行者,根据其职

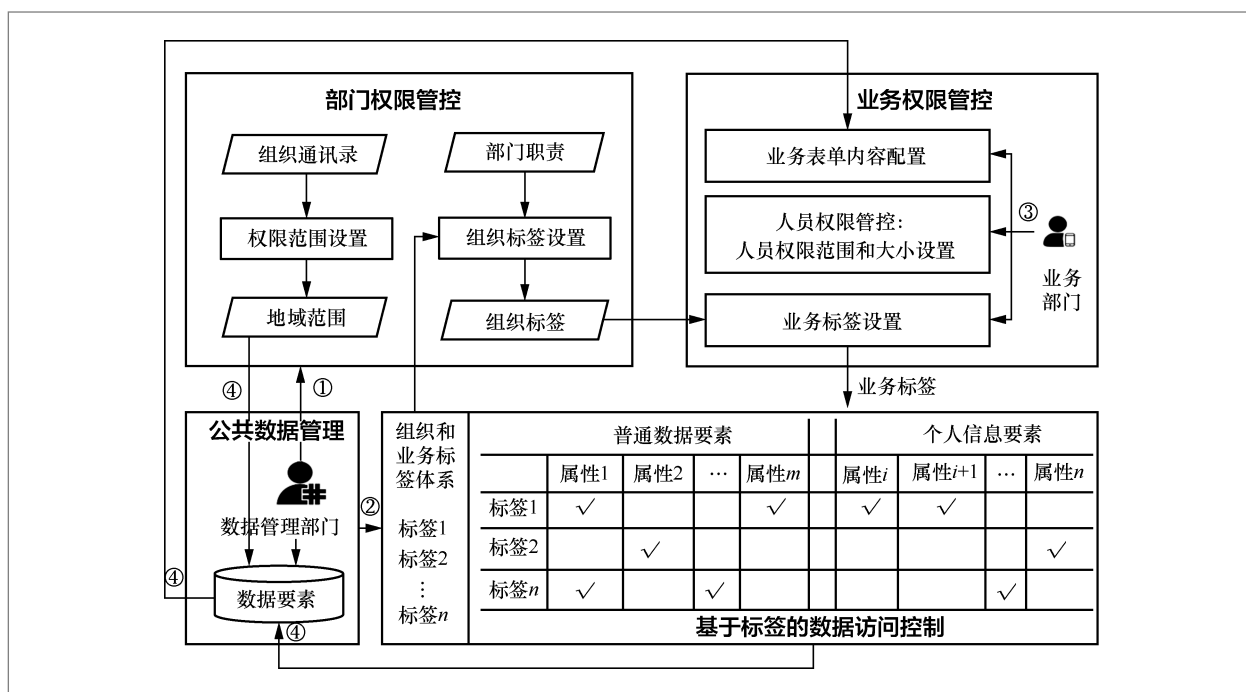


图5 政府侧的数据授权技术体系

责生成对应的标签。一个业务的性质由若干个业务标签表示,是该部门组织标签的子集,由业务部门根据具体业务需要从本组织拥有的标签中选定。例如,民政局的组织标签包含老人关怀、婚姻登记、殡仪服务、社会救助、养老陪护、儿童收养等。民政局在一项特困老人慰问业务中,可选择“老人关怀”和“社会救助”两个标签来表示业务性质。

● **业务权限管控模块:**如图5流程③所示,业务部门在利用信息系统配置一项具体业务时,可配置需采集的信息表单内容,选择该业务关联的标签,并设定基层工作者在此项业务中的权限大小。例如,市民政局希望开展独居老人关怀慰问工作,可基于政民互通平台^[7]提供的台账功能,利用其中的数字化控件^[5]配置承载该业务的“社区老人台账”。在台账配置过程中,可从民政局关联的组织标签中选择“老人关怀”和“福利发放”两个标签来表明该业务的性质。进一步,业务部门在台账模板中设置基层工作者的权限内容和大小,如能否共享和转让台账、能否导出业务管理数据、能否代居民填报等。此后,基层工作者基于民政局配置的台账模板,分别申领各自辖区内的“社区老人台账”,并按照台账模板设定的权限大小来管理和使用台账中的数据。

● **数据访问控制模块:**利用基于标签的访问控制技术,数据管理部门为每个标签设定可使用的普通数据要素和个人信息要素中的属性(城市知识体系中的属性,而非业务字段),精确管控部门和业务的数据权限。同时,该模块也会根据一个部门管辖的地域范围来控制其能获取的数据内容。如图5流程④所示,根据一个业务设定的标签和一个部门的管辖范围,从数据要素中获取对应地域、对应要素属性的数据,回填至业务表单中,以满

足业务需求。

(2) 用户个人授权管理技术

图6展示了用户个人授权管理部分的技术实现方式,详细机制在第2.4节的第二部分进行了描述。

用户在应用A和B中填报的个人信息会通过数字化控件进入业务数据库。之后,智能系统(如文献[5]中提出的数字化控件和城市知识体系)可自动判断这些个人信息是否已经包含在个人信息要素中。当有新的个人信息产生时,系统通过逐个提示的方式询问用户是否将该信息添加至“我的”个人信息中。如用户选择“确定”,这些个人信息被写入底层的个人信息要素库,并可在“我的”界面中进行查看和编辑。

用户也可以主动在“我的”界面添加个人信息,系统会一一询问是否进行个人信息的授权(除非用户明确表达“今后不再重复提示”)。一旦用户在“我的”界面对某条个人信息进行删除操作,对应的底层个人信息要素也会被随之删除。因此,个人信息要素的数据资源持有者永远被用户个人掌握。

如果个人信息要素是通过智能检测提示的方式从其他服务中获取的,用户在“我的”界面删除此条个人数据要素后,对应的底层个人信息要素会被随之删除,但业务数据库中对应的个人信息作为公共服务业务数据仍然存在(仅限于在原业务中使用,不可跨服务使用)。用户如果希望进一步删除原业务中的个人信息,可在对应的服务中删除相关记录。因此,个人信息数据的数据资源持有者也在用户手中。

3 结束语

本文将公共数据划分为两大类四小类:公共业务数据(包括业务管理数据和

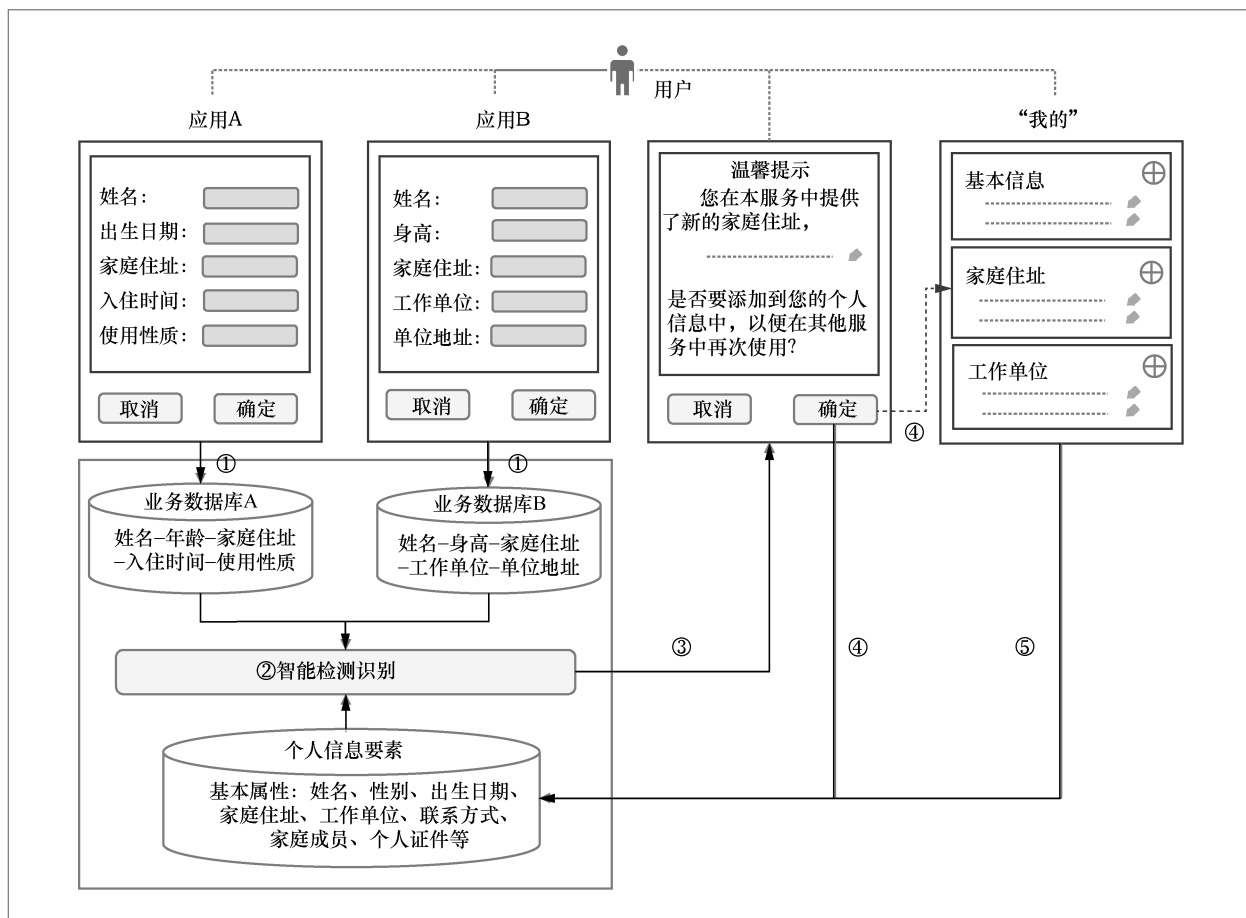


图6 个人信息授权管理的技术实现方法

个人信息数据)和公共数据要素(包括普通数据要素和个人信息要素),针对公共数据要素的4个重要参与方(个人、基层工作者、数据管理部门和政府业务部门),提出了数据资源持有权、数据加工使用权和数据产品经营权分置的公共数据产权制度,为4类公共数据设计了详细的数据确权 and 授权机制,并研发了配套的技术方案,可解决因基层工作者数据权限不明导致的用数难问题、因部门责权不清晰带来的工作瓶颈问题、因个人信息授权模糊带来的安全和效率问题,从而促进公共数据的安全、合规、有序流通,并可为公共数据在更大范围进行循环打下基础。该方案已经在北京开展试点,达到了预期

的效果。文献[6]、文献[4]和本文组成了公共数据要素三部曲,分别从业务架构、底层技术和运行机制3个方面支撑公共数据要素的构建。

参考文献:

- [1] 朱扬勇. 依照数据用途界定公共数据[J]. 大数据, 2024, 10(3): 163-167.
ZHU Y Y. On public data[J]. Big Data Research, 2024, 10(3): 163-167.
- [2] 梅宏, 杜小勇, 金海, 等. 大数据技术前瞻[J]. 大数据, 2023, 9(1): 1-20.
MEI H, DU X Y, JIN H, et al. Big data technologies forward-looking[J]. Big Data Research, 2023, 9(1): 1-20.

- [3] 杜小勇, 陈跃国, 范举, 等. 数据整理: 大数据治理的关键技术[J]. 大数据, 2019, 5(3): 13-22.
DU X Y, CHEN Y G, FAN J, et al. Data wrangling: a key technique of data governance[J]. Big Data Research, 2019, 5(3): 13-22.
- [4] 郑宇, 易修文, 齐德康, 等. 基于城市知识体系的公共数据要素构建方法[J]. 大数据, 2024, 10(4): 130-148.
ZHENG Y, YI X W, QI D K, PAN Z Y. Elementarisation method for public data based on urban knowledge systems[J]. Big Data Research, 2024, 10(4): 130-148.
- [5] 梅宏. 数据治理之法[M]. 北京: 中国人民大学出版社, 2022.
MEI H. Methods of data governance[M]. Beijing: China Renmin University Press, 2022.
- [6] 郑宇. 政民互通: 构建政府和居民之间的双向信息通道[J]. 大数据, 2024, 10(1): 127-140.
ZHENG Y. Building bidirectional digital channels between governments and citizens[J]. Big Data Research, 2024, 10(1): 127-140.
- [7] 郑宇. 城市治理一网统管[J]. 武汉大学学报(信息科学版), 2022, 47(1): 19-25.
ZHENG Y. Unified urban governance models[J]. Geomatics and Information Science of Wuhan University, 2022, 47(1): 19-25.
- [8] 张峰, 李照川, 王冠军, 等. 公共数据确权的探索与实践[J]. 信息通信技术与政策, 2023, 49(4): 34-41.
ZHANG F, LI Z C, WANG G J, et al. An exploration and practice of public data property rights confirmation[J]. Information and Communications Technology and Policy, 2023, 49(4): 34-41.
- [9] 张会平, 薛玉玉. 公共数据授权运营产权运行机制的理论建构与实施路径[J]. 电子政务, 2023(11): 1-13.
ZHANG H P, XUE Y Y. Theoretical construction and implementation path of property right operation mechanism of public data authorization operation[J]. E-Government, 2023(11): 1-13.
- [10] 凌捷. 大数据安全与隐私计算[J]. 大数据, 2023, 9(6): 1-2.
LING J. Big Data Security and Privacy Computing[J]. Big Data Research, 2023, 9(6): 1-2.
- [11] 杜小勇, 柴云鹏. 跨域数据管理[J]. 大数据, 2023, 9(4): 1-2.
DU X Y, CHAI Y P. Cross-Domain Data Management[J]. Big Data Research, 2023, 9(4): 1-2.

作者简介



郑宇(1979-),男,博士,京东集团副总裁、京东智能城市研究院院长、京东科技首席数据科学家,IEEE Fellow、美国计算机学会杰出科学家、KDD China主席,国家“万人计划”科技创新领军人才,享受国务院政府特殊津贴专家,上海交通大学讲座教授,南京大学、香港科技大学等多所高校客座教授。先后担任人工智能顶尖国际期刊*ACM TIST*的主编、国家重点研发计划项目首席科学家及总负责人,以及ICDE、CIKM等多个国际会议的程序委员会主席。提出“城市计算”理论体系,主导的智能城市操作系统成为雄安新区的数字底座,经国家批复,雄安新区的智能中枢以“雄安城市计算中心”命名。

收稿日期: 2024-06-05

基金项目: 国家自然科学基金项目(No.62076191);北京市科技计划项目(No.Z211100004121008)

Foundation Items: The National Natural Science Foundation of China(No. 62076191), Beijing Science and Technology Plan(No. Z211100004121008)