**深圳市地方标准**

**智慧城市数据账户资源共享应用规范**

**Application specification for resource sharing of data account for smart city**

**（征求意见稿）**

**编 制 说 明**

2021-6

**一、任务来源**

根据深圳市市场监督管理局于2020年5月12日下达的《深圳市市场监督管理局关于下达2020年第一批深圳市地方标准计划项目任务的通知》，《智慧城市数据账户资源共享应用规范》标准计划编号为98号。计划完成日期为2022年6月。

本文件由深圳市政务服务数据管理局归口。
本文件建议作为推荐性深圳市地方标准。

**二、编制背景、目的和意义**

**（一）编制背景**

1. 应用背景

我国智慧城市经十多年的历程得到了长足的发展，数字政府初露端倪，而数据是构筑智慧城市及数字政府的核心内容。但由于数据相关部门的权责及业务划分划分、数据工具存量不足和标准滞后等原因，导致高效开展智慧城市的相关应用、业务及服务仍具有一定的困难。

首先，各部门的数据均针对各自领域的业务和服务进行采集、存储、整理及管理，部门内部建有数据库或数据管理平台，不同部门之间数据库或数据管理平台的交互在数据内容、数据架构、数据模型、数据标准、数据质量等方面不对称，导致跨层级、跨系统、跨部门、跨业务的数据资源难以便捷、顺畅的流通和应用，对复杂综合业务和精准创新业务的支持力度难以充分发挥。

其次，目前大部分地区各部门的数据交换仍停留在“数据共享交换平台”模式上，业务场景中所需数据资源难以从众多数据源中看清楚，数据服务申请仍需通过手动开发模式生成服务接口，服务申请周期较长，接口缺乏灵活性，无法支持数据安全脱敏机制，同时也难以避免接口重复开发的问题，从而导致数据共享开放效率低、成本高、安全性也难以得到有效保障。

再次，数据共享交换平台中不同数源数据重复度高，数据标准和数据质量差异较大，数据来源难以汇总归集，数据问题追踪困难、数据更新校正周期长，导致数据服务和业务应用难以统筹规划、综合应用，更无法实现业务的反向治理和数据的反向补充。

其四，任何单一部门的数据往往只能反映数据主体的一个侧面，用户全生命周期的数据散落在众多业务系统的资源中，现有的数据管理平台还未做到对众多业务线和众多数源中的用户数据进行提炼、加工，从而进行全貌的展现和汇集，更无法从分散的数据中发现主体对象的数据特征和用户画像，挖掘数据中潜在的业务价值，提高数据共享、流通效率，加速数据资产化应用成效。

其五，传统数据治理模式更偏重于技术侧的元数据、主数据的管理、数据实施过程中的质量问题发现和处理以及数据安全审计等方面考虑，在业务测的治理规划、业务资源下探、数据质量改进以及数据流通闭环上的考虑较为薄弱，导致数据管理混乱，缺少清晰明确的体系化结构和治理流程监督跟踪机制，从而难以解决长期存在的数据覆盖度较低，资源丰富度不足，服务流动性不够，传统的数据治理模式与高阶的数据资源化、数据资产化无法更有效的匹配。

1. 政策法规背景

近几年，数据相关政策密集出台，数据标准被当作实施智慧城市及数字政府国家战略的重点任务。

2019年2月，中共中央、国务院《粤港澳大湾区发展规划纲要》提出建成智慧城市群，推进新型智慧城市试点示范和珠三角国家大数据综合试验区建设，加强粤港澳智慧城市合作，探索建立统一标准，开放数据端口，建设互通的公共应用平台；2019年8月，中共中央、国务院《关于支持深圳建设中国特色社会主义先行示范区的意见》提出加快建设智慧城市，支持深圳建设粤港澳大湾区大数据中心；2020年10月，《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》提出加强数字社会、数字政府建设，建立数据资源产权、交易流通、跨境传输和安全保护等基础制度和标准规范；2021年1月，深圳《关于加快智慧城市和数字政府建设的若干意见》提出到2025年，打造具有深度学习能力的鹏城智能体，成为全球新型智慧城市标杆和“数字中国”城市典范，建设城市数字底座，打造城市智能中枢，推进业务一体化融合，实现全域感知、全网协同和全场景智慧，让城市能感知、会思考、可进化、有温度；2021年2月，广东省提出建立数据资源基础制度和标准规范，支撑加快数字经济、数字社会、数字政府发展。

第十三届全国人大常委会第二十八次会议对《中华人民共和国数据安全法（草案二次审议稿）》进行了审议，该法提出“规范数据处理活动、保障数据安全、促进数据开发利用”，从立法角度赋予数据新的责任与内涵。

**（二）编制目的**

首先，本标准通过构建“数据账户”，归集个人与企业两大“主题”全生命周期数据，根据应用、业务及服务需求，建立数据账户目录，并预留扩展接口，将多部门采集的数据进行融合，覆盖主题的全生命周期，解决数据不全面的问题。

其次，本标准提出建立统一标准化的服务接口，提供订阅、查询、推荐的应用内容，并为敏感信息建立脱敏规则，解决服务接口重复手动开发的低效、高成本与共享应用的安全性问题。

再次，本标准提出“一数一源”的要求，与深圳市公共基础信息的数据来源要求保持一致，对数据来源进行唯一性界定，打破数据孤岛，解决数据分散不成体系问题。

其四，本标准针对归集融合后的数据提出特征标签、特征画像的高阶数据应用形态，为共享应用方式从支持传统的查询模式，升级为精准“数据对接”模式，支持精准的“智能提醒”、“智能秒批”、“轻松秒填”、“刷脸秒办”、“无感申办”等共享应用。

其五，本标准提出“资源治理”的数据治理新模式，将面向基础数据的传统数据治理进行面向业务数据的资源治理延伸与分层优化，形成逻辑更加清晰的大数据治理结构，数据与业务的结合更紧密，使治理效率更高、效果更好，面向业务应用提供更精准的大数据支撑服务。

**（三）编制的意义或必要性**

1. 优化政务数据管理

打通信息壁垒，实现跨部门、跨层级系统的业务融合和数据集约；推动建立“一数一源”管理机制，解决政务数据的精准化；构建数据账户，形成最权威、最全面的个人与企业全生命周期的信息账户。

1. 提升政务服务效能

实现“一（网）厅”办理所有事项；提供“秒填、秒证、秒批、秒办”的数据应用服务，促进我国电子政务高质量发展；为政府决策科学化、社会治理精准化、公共服务高效化提供科学指导。

1. 节约行政资源成本

节约投入成本，避免系统重复建设、数据重复处理、接口反复开发、数据多次调用。

1. 社会价值

促进城市事务更加文明有序。数据账户平台数据资源面向政务应用全面开放、面向社会合理适度开放，推动社会公共服务优化升级，提升民生幸福体验。

促进政府服务优化升级。助力于优化政府治理结构，推动资源服务走向智能化、智慧化，加速简政放权的步伐，推动政务服务走上更高标准的新台阶。

用数据治理带动社会的治理走向新高度。完善数据治理，建立面向业务过程、服务过程的数据资源治理，反向助力改善政务服务流程与数据体系。

1. 经济价值

提高资源服务效率。传统模式下，资源供给方在提供资源服务时，通常需要改造、重构甚至重建系统，投入巨大但收效有限。数据账户借用新技术、新手段安全高效的实现资源的服务化，极大的提升了资源的使用效率和资源服务的覆盖面。

降低资源获取成本。数据账户通过数据服务化，提供标准化的数据应用工具 ，有效改变现有需要花费大量人力物力财力寻找、获取资源的现状，让大数据应用领域的各相关方极大的减少资源的获取成本与使用成本。

实现技术与创新双引领。数据账户是深圳市在数据共享应用和隐私安全保护等方面做出的大胆尝试和创新探索，综合运用大数据、区块链、云计算、AI等技术，解决政府服务于民众最后一公里的问题。

**三、主要起草过程**

本标准从2019年预研开始，经过项目调研、标准研究，于2020年5月立项，经过标准编制组的努力，多次召开标准讨论会，编制组内部达成统一，形成了目前的征求意见稿。

2019年，结合数据账户在龙华区的应用，开展相关的数据账户建设、运行情况调研，并同步开展政务数据、政务信息资源交换体系、政务信息资源目录体系、电子政务、政务数据开放共享等标准研究。

2020年1月-3月，进行立项文件的编写及行业行政主管部门申报工作。

2020年3月-5月，经过行业行政主管部门和标准化主管部门对立项文件的形式审查、技术审查，正式下达立项计划。

2020年5月-8月，成立由数据账户管理方、建设方、需求方相关的部门、企业及机构的专家组成的标准编制组，进行调研及相关案例的征集。

2020年9月-11月，编制组通过对数据账户相关概念的定义、在智慧城市中的定位和应用边界、资源及要求、建设、共享应用以及管理要求的研究，相关资料的整理，标准文本的编写，形成标准的工作组讨论稿。

2020年11月-2021年3月，召开了标准的启动会和多次标准讨论会，针对工作组讨论稿进行讨论、修改。

2021年3月，与深圳市大数据资源管理中心进行标准的技术内容研讨，明确了本标准与《公共基础信息数据元规范》的协调关系。

2021年4月，编制组针对标准工作组讨论稿进行了2次讨论及修改。

2021年5月-6月，编制组完善标准内容并形成标准征求意见稿。

**四、制定标准的原则和依据，与现行法律、法规、标准的关系**

编制组遵循“科学性、统一性、规范性、连续性、一致性”的原则，在编制过程中严格按照我国现行有效的国家标准和行业标准的要求，引用文件准确合理，文本结构严谨、逻辑清晰。

本标准制定过程中引用的主要标准如下：

GB/T 21062.2 政务信息资源交换体系 第2部分：技术要求

GB/T 21062.3 政务信息资源交换体系 第3部分：数据接口规范

GB/T 22081 信息技术 安全技术 信息安全控制实践指南

GB/T 25070 信息安全技术 网络安全等级保护安全设计技术要求

GB/T 32419.4 信息技术 SOA技术实现规范 第4部分：基于发布/订阅的数据服务接口

GB/T 32908 非结构化数据访问接口规范

GB/T 34960.5 信息技术服务 治理 第5部分：数据治理规范

GB/T 37722 信息技术 大数据存储与处理系统功能要求

GB/T 38664.2 信息技术 大数据 政务数据开放共享 第2部分：基本要求

DB4403/T XXX.1 公共基础信息数据元规范 第1部分：总则

DB4403/T XXX.2 公共基础信息数据元规范 第2部分：人口

DB4403/T XXX.3 公共基础信息数据元规范 第3部分：法人

《政务信息资源目录编制指南（试行）》

**五、主要参加单位和工作组成员及其所做的工作等**

表1 主要参加单位和工作组成员及其所做的工作

|  |  |  |
| --- | --- | --- |
| **工作单位** | **姓名** | **项目分工** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**六、主要技术内容及技术依据**

**（一）标准的编制思路**

本文件以“明确数据账户在智慧城市应用中的定位以及边界—界定数据账户的资源及要求、数据账户资源的数据来源及要求—提出数据账户的建设—数据账户如何提供共享应用—数据账户管理要求”的逻辑顺序进行标准的技术内容编制。

**（二）标准的范围**

本文件规定了数据账户在智慧城市中的应用框架模型、资源、数据来源、建设、共享应用和管理要求。

本文件适用于在智慧城市中以个人与企业为主题进行数据账户的设计、建设、运行及应用。其他主题的数据账户可参照使用。

**（三）术语和定义**

定义了本体、值数据、明细数据、本体指标、特征标签、特征画像、数据账户、个人数据账户、企业数据账户、扩展数据账户、资源治理、目录资源、数据资源、数据账户资源、通用服务、自定义服务、数据来源方、数据需求方、数据管理方等术语。

本体

把智慧城市中可进行数据和信息归集的对象作为数据账户中的“本体”，所归集的“数据”和“信息”具有相关的“属性”及“关系”的特征。

“本体”概念的提出旨在提出一种事务的“本源”属性，事务就是所涉及的对象，“本源”代表自身的属性，在计算机或网络中可用数据及信息来表示。

值数据

数据账户中具有单一数值的数据，定义为值数据。

明细数据

数据账户中包含详细属性信息的一系列数据，用“详细属性信息”和“一系列”进行“明细”的定义。

本体指标

本体的基础数据属性信息，数据账户中的值数据和明细数据表示本体指标的具体内容。

特征标签

根据业务需求所设定的规则，基于归集的数据资源，由数据账户系统形成的对本体指标的一类新的特性标示。

特征画像

基于特征标签，形成的更高层次和全貌的特性标示。

数据账户

“账户”是一类数据和信息的集合，本文件是通过个人和企业两大本体建立数据账户，抽象成为“一定主题分类方式”进行归集，并覆盖本体的全生命周期过程中涉及的数据和信息，数据账户中本体可以是智慧城市中的任何本体对象。

个人数据账户

以个人为本体建立的数据账户。

企业数据账户

以企业为本体建立的数据账户。

扩展数据账户

以个人和企业之外的其他本体建立的数据账户。

资源治理

区别于数据治理，将数据治理后的数据，进行资源治理，形成面向业务的资源，提供资源服务。资源治理的过程是先根据业务需求建立目录和本体指标，再将数据治理后的数据按目录和本体指标进行归集融合，形成数据账户资源。

目录资源

通过本体建立的目录及本体指标，定义为目录资源。

数据资源

基于目录资源的需求，从数据源中采集到的数据并经过数据治理和资源治理后的数据定义为数据资源。

数据账户资源

经过资源治理后的目录和数据，目录资源和数据资源统称为数据账户的资源。

通用服务

在数据账户系统中已按业务需求及相关规则明确并提供的数据服务。

自定义服务

区别于通用服务，可由数据账户的应用相关方按业务需求进行自主申请的数据服务。

数据来源方

数据账户获取数据的部门，例如各相关的职能部门等，在自定义服务的情况下，是应用数据账户的个人或企业。

数据需求方

应用及使用数据的相关方，如个人或企业。

数据管理方

应用数据账户提供服务，并对相关数据进行管理的部门，例如职能部门。

**（四）数据账户应用框架模型**

数据账户定位于将从公共基础信息、数据共享交换平台或其它业务系统来源的数据进行数据治理和资源治理，形成数据账户的资源，通过服务接口的方式为智慧城市的相关业务提供共享应用。

**（五）数据账户资源**

根据定义，数据账户资源包括目录资源和数据资源。目录资源是通过业务的需求，围绕本体进行建立，并形成目录资源的层级管理，可进行目录层级的无限扩展，以支持根据更多的本体建立数据账户并应用。数据资源应具有可信的数据来源，并与数据来源保持数据版本的一致性，才能满足对业务的支撑。

数据账户资源来源于三个部分，第一个是深圳市的公共基础信息，依据《公共基础信息数据元规范》，第二个是各部门的数据共享交换平台，第三个是其他的业务系统，为更多的第三方应用进行功能预留。

从三个部分来源的数据，首先要进行数据治理，然后进行资源治理，形成数据账户的数据资源，通过资源服务对外提供共享应用。目录资源需要按业务需求以及主题进行建立，按目录层级进行管理，并能无限扩展，目录的最后一级为本体指标，需要对本体指标的数据类型进行明确定义。个人数据账户的本体指标符合《公共基础信息数据元规范》系列标准中人口部分的数据元和数据来源要求，以标准中公民身份号码为唯一标识，且应覆盖个人的生命周期各环节。企业数据账户的本体指标符合《公共基础信息数据元规范》系列标准中法人部分的数据元和数据来源要求，以标准中统一社会信用代码为唯一标识，且应覆盖企业的生命周期个环节。扩展数据账户目录按业务需求进行主题目录的建立，并以本体的唯一标识符进行标识，覆盖本体的全生命周期各环节。数据资源的数据来源于公共基础信息、数据共享交换平台和其它业务系统，数据来源应可信，并与数据来源保持更新的一致性和相同的数据版本，保证基础数据的动态更新。

数据作为数据账户应用的核心，从数据质量以及数据来源方就需要满足相应的要求。对于数据主要从格式规范、内容完整进行规定，对于数据来源方，本体指标应具有唯一的数据来源，才能保障为业务应用提供精准的数据支撑，数据来源方应对数据进行有效性验证才能满足数据账户对所归集数据的需求。对于数据管理方，应按业务需求建立数据账户的目录和本体指标，并保证本体指标数据来源的唯一性，并对经过数据治理后的数据进行资源治理，形成数据资源，为面向业务的共享应用服务提供数据资源的支撑。

**（六）数据账户建设**

数据账户的建立由数据目录体系、数据接入体系、数据存储体系、数据运算体系、数据服务体系、数据管理体系等技术体系组成，为数据账户的目录资源、数据资源和资源服务的建立提供技术体系支撑。目录资源通过数据账户的数据目录体系构建，完成目录管理和本体指标建立；数据资源由数据接入体系、数据存储体系、数据运算体系构建，完成数据资源采集、打标和画像的过程。资源服务将数据资源通过数据服务体系提供共享应用；数据管理体系提供对数据账户其他技术体系的管理。

数据目录体系

数据目录体系是数据账户的核心，根据应用场景和业务需求建立一级目录、二级目录、扩展目录、本体指标的目录资源，根据目录资源从公共基础信息、数据共享交换平台或其他业务系统归集数据。

数据目录体系应具有根据业务对本体指标进行重组的功能，并在本体指标的来源无法采集时补充符合要求的可信数据来源，在目录内已经确定的目录层级结构和内容由数据管理方根据业务需要进行修改，同时为扩展至第三方应用，需要支持数据的自定义功能，例如个人和企业可对需要增补或调整的内容进行自定义调整，但自定义的数据仅能作为个人自行使用或企业内部使用，无法经由数据账户系统进行共享应用，需要进入数据账户进行共享应用的，应经过相关的认证或许可以保证数据质量和安全性等要求，数据账户内的目录和数据条目通过编码的方式保证唯一性和可追溯性，为扩展到更多的应用，数据目录体系应支持账户的无限扩展。

数据接入体系

从公共基础信息、数据共享交换平台或其他业务系统归集到的数据，通过数据接入体系的抽取、清洗、转换、加载后，需要满足时效性、一致性、规范性、可访问性等数据质量的要求才能达到数据账户的使用需求。

数据存储体系

针对数据账户中的海量数据，为保证安全、可用、高效的数据存、取能力，建议采用分布式存储方式，具有数据分布式写入、数据分布式存储、数据分布式访问和数据存储管理等功能模块，并满足相应的功能要求，支撑数据账户对数据存储的要求。

数据运算体系

数据运算体系包含计算引擎、算法模型库、计算资源、数据的输入和结果输出等功能，并能支持数据账户的通用服务和自定义服务。计算引擎需要能支持海量账户数据的实时计算和离线计算，算法模型库能支持算法模型的动态扩展，数据输入来源于数据存储体系，结果输出能够将数据写入存储体系，也能够提供给数据服务体系。

数据服务体系

数据服务体系在数据账户内部需要支持目录资源查询、数据资源查询、数据过滤、数据匹配、数据比对等服务，对数据需求方提供服务时应具有数据服务接口、数据服务调用、数据服务监控、数据服务分析等功能。数据服务接口符合GB/T 32419.4中的规定，数据服务的调用需要具有不同的权限，数据服务监控应定期对数据服务接口使用情况和服务运行情况进行监控，对监控所采集到的信息还需要进行相关的分析，为数据服务的管理提供参考和依据。

数据管理体系

数据管理体系是对整个数据账户技术体系的管理，具有账户目录管理、账户资源管理、账户数据治理、数据服务管理、数据分析管理等功能。

账户目录管理需要维护并管理数据账户目录资源，能对账户目录进行编制、发布，并提供共享策略和队策略的更新；数据资源的管理，需要进行数据查询、数据纠偏申请，管理数据来源、数据接入，对数据进行更新，对数据量进行管理；账户数据治理；对账户资源进行全生命周期的数据治理，需要对数据标准、数据模型、元数据、主数据、数据质量、数据安全等进行管理；对数据服务进行管理，需要进行数据服务查询、数据服务订阅、数据服务审核、数据服务统计、数据服务授权、数据服务认证、数据地图等的管理；数据分析管理需要管理根据数据资源运算结果提供的加工数据服务，并能够对本体指标、特征标签以及特征画像的分析功能进行管理。

**（七）共享应用**

数据账户的共享应用通过服务接口的方式提供，并按“服务上架-服务申请服务调研-服务监控”的流程进行，支持订阅、查询、推荐以及其他扩展应用场景。

订阅服务是数据需求方发起申请，数据管理方审核通过后提供应用服务，查询是数据需求方以模糊查询和精准查询的方式对数据账户发起查询，数据账户系统按条件反馈结果的方式，推荐是由数据需求方提出推荐条件，数据账户匹配条件后将结果推荐给数据需求方。

数据账户以服务接口的方式提供服务，接口符合相应的标准即可。

服务上架是数据账户将服务接口封装成不同的服务发布，数据需求方按服务申请，服务申请是数据需求方按应用需求申请已上架的服务，如上架服务中没有需要的服务，则发起服务订阅申请，服务调用是数据需求方的服务申请经过数据管理方的权限验证后由数据账户提供服务调用，服务监控是数据管理方在服务调用过程中对服务接口的使用进行监控。

**（八）管理要求**

主要对账户资源编码、元数据和安全进行管理。

账户资源编码应具有唯一性，以便于数据和信息的定位、追溯、管理以及精准服务的提供；元数据的管理符合国家已有标准的相关规定，对于公共基础信息，符合《公共基础信息数据元规范》中的规定；安全管理主要是网络和数据的安全，符合国家标准已有的相关规定；对于敏感信息，数据账户本体指标中的敏感信息可参照YD∕T 3473进行内容定义及分类，并根据应用需求进行脱敏等级划分和建立相应的脱敏规则。由附录给出个人数据账户和企业数据账户本体指标中的敏感信息脱敏规则以及脱敏等级参考。

**（九）附录**

由于实际应用过程中目录资源较多，在资料性附录A中提供目录资源的示例，本体指标覆盖《公共基础信息数据元规范》中人口与法人部分，扩展部分以及敏感信息脱敏等级供参考使用；资料性附录B中提供数据账户中本体指标的值数据、明细数据应用示例，附录C中提供共享应用的流程、提供方式以及接口的应用示例，附录D给出本体指标敏感信息的脱敏规则。

**七、明确标准中涉及专利的情况，对于涉及专利的标准项目，应提供全部专利所有权人的专利许可声明和专利披露声明**

本文件不涉及知识产权问题。

**八、重大意见分歧的处理依据和结果**

无重大意见分歧。

**九、标准的属性**

本标准为深圳市地方标准。

**十、实施标准的措施建议**

标准发布实施后，行业主管部门应向相关单位及时通报标准发布信息，做好宣传和培训。

**十一、其他说明**

无。

 《智慧城市数据账户资源共享应用规范》标准编制组

 2021年6月