

# 面向智慧城市的大数据开放共享平台的设计与实现

孙傲冰<sup>1, 2</sup> 季统凯<sup>1, 2</sup> 伍小强<sup>3</sup>

(1. 国云科技股份有限公司 广东 东莞 523808; 2. 东莞中国科学院云计算产业技术创新与育成中心 广东 东莞 523808;

3. 广东电子工业研究院有限公司 广东 东莞 523808)

(\* 通信作者电子邮箱 sunab@casc.ac.cn)

**摘 要:** 针对建设中数据规划及管理存在的问题, 提出一种从政务数据资产的角度入手定义智慧城市大数据空间模型, 对智慧城市大数据开放共享平台进行总体设计的方法。通过合理划分空间维度及空间抛面, 建立大数据模型、业务系统、数据实例、及数据目录的映射关系, 实现智慧城市统一数据库的建设。建立大数据安全访问控制模型, 通过大数据统一访问接口, 将加工后的数据供给授权的第三方应用开发商调用。按照“一片云、一张网、一套库、一把锁、一张图”的思路在东莞搭建智慧城市大数据开放共享平台原型系统, 基于城市用户的庞大用户群, 建设一个应用发布、数据定制及评价的统一门户。通过实例验证平台可以实现数据提供者、应用开发者及用户相互促进的智慧城市数据及应用的共建共享, 打造需求、数据、资金合理流向的产业生态。

**关键词:** 大数据; 智慧城市; 数据资产; 数据空间; 开放共享平台

**中图分类号:** TP311.132 **文献标志码:** A

## Design and realization of big data open platform for smart city

SUN Aobing<sup>1, 2</sup>, JI Tongkai<sup>1, 2</sup>, WU Xiaoqiang<sup>3</sup>

(1. G-Cloud Science and Technology Limited Company, Dongguan Guangdong 523808, China; ;

2. Cloud Computing Center, Chinese Academy of Sciences, Dongguan Guangdong 523808, China;

3. Guangdong Electronic Industrial Institute, Dongguan Guangdong 523808, China)

**Abstract:** Due to the planning and management problems in traditional smart city construction, a big data space model for smart city from the point of data asset was proposed. Through the division of spacial dimension and parabolic surface, the mapping relations among big data model, business system, data instance and data directory were created for the unified big database construction. The big data security access control model was established for data extraction and data cleaning under different scenes. Through the unified cross-language platform interface, the big data could be accessed by the third party application developers. According to the idea of “a cloud, a net, a library, a lock, a map” for smart city, one instance at Dongguan was built for big data open platform, and one portal for smart city for data or applications customization and evaluation was developed for data providers, application developers and users. The industrial ecology was made with reasonable flow of demand, data and capital based on smart city's huge users.

**Key words:** big data; smart city; data asset; data space; open platform

## 0 引言

2013 年以来欧美信息技术强国都在推动智慧城市建设与大数据技术的结合, 以面向市民提供更加精准、实时、个性化的社会服务<sup>[1]</sup>。2013 年 6 月, 日本公布了“创建最尖端 IT 国家宣言”, 提出开放公共的政府及社会服务数据, 以大数据技术为核心把日本建设成为一个具有“世界最高水准的广泛运用信息技术的社会”。2014 年 3 月, 美国政府整合 6 个部门投资 2 亿美元启动“大数据研究和发展计划”, 推动大数据应用与社会服务的紧密结合。2015 年 6 月, 英国宣布将在政府信息化基础设施方面投入巨资, 加强围绕社会服务数据采集和分析<sup>[1-3]</sup>。我国各级地方政府也积极探索利用大数据技术, 促进政府的数据流通, 推动智慧城市的实现<sup>[4]</sup>。如上海市政府数据资源向社会开放工作, 在 2015 年已确定 190 项数据内容作为 2014 年重点开放领域。从市场角度来看, 企业主

体是嗅觉最灵敏, 最能挖掘用户需求, 并提供最佳服务的主体。如 2012 年开始, 国家民航信息中心也开放其数据运营接口, 从而在此基础上开发了众多的面向航空服务的移动端 APP 软件。

对于政府而言, 数据在各种智慧城市系统中“睡眠”是毫无价值的, 只有让数据流动起来, 在政府监管的前提下让市场决定其去向, 才能充分发挥其价值, 从而为群众提供更加优质的服务<sup>[3-4]</sup>。2014 年以来我国密集制定了《关于促进智慧城市健康发展的指导意见》《互联网+行动计划》《运用大数据加强对市场主体服务和监管的若干意见》《促进大数据发展行动纲要》, 推进数据开放共享, 推动智慧城市的建设。

## 1 现状及分析

自 20 世纪 90 年代以来, 我国各级政府部门围绕各自业务系统, 开发了为数据众多的业务系统。各种数据如只在单

收稿日期: 2016-09-23; 修回日期: 2016-10-20。 基金项目: 国家住建部智慧城市科技研发项目(2016-K3-008); 广东省科技厅资金项目(201001D0104726115, 2014B010118001, 2015B010131001, 2015B010109001)。

作者简介: 孙傲冰(1978—), 男, 河南新乡人, 副研究员, 博士, CCF 会员, 主要研究方向: 云计算、物联网、大数据; 季统凯(1972—), 男, 吉林长春人, 研究员, 博士, CCF 会员, 主要研究方向: 云计算、物联网、大数据; 伍小强(1964—), 男, 广东珠海人, 高级工程师, 博士, 主要研究方向: 大数据、图像处理。

一的系统内流动,其价值只发挥了 10% 不到。智慧城市数据只有突破单一系统的界限,从在一个系统内流动到被尽可能多的系统共享,才能实现价值的最大化<sup>[5]</sup>。我国各级政府部门中仍存在一些“意识壁垒”,困扰了现有的智慧城市系统及现有电子政务系统间的数据共享:

一是数据所有权问题:一些政府部门认为智慧城市业务系统是哪个部门管理的,系统中的数据就是哪一个部门的,把数据共享出去将会造成其利益的损失<sup>[4,6]</sup>。

二是数据安全责任问题:因为智慧城市安全的管理归其管理部门,一些部门认为只有不共享才能保证不泄密。随着当前信息技术的进步,数据在安全授信条件下,实现共享的技术已经成熟,只有数据共享才能为用户提供更多便利的服务,智慧城市数据共享也是大势所趋。

三是数据价值认知问题:一些部门认为其管理的系统中的数据只对其业务有价值,没有认识到其数据对于其他部门或对于企业和市民提供服务的价值。

四是数据精确性认知的问题:一些智慧城市的信息系统缺乏统一管理的流程和规范,导致采集到的数据质量不高,不同的时间、地点或操作人员采集的数据质量不一致,导致数据间相互矛盾,存在较多的“数据冲突”。

五是数据采集意识局限性的问题:一些与智慧城市业务相关的信息系统沿用了多年,采集、汇总和分析手段未能随信息技术的发展共同进步,如还在使用老的单机数据库系统,导致数据采集的效率不高,无法有效共享<sup>[4,7]</sup>。

## 2 智慧城市大数据模型

### 2.1 设计依据

定义 1 数据资产是指企业、自然人、国家拥有或者控制的能以价值或有用性来计量或判断的数据资源的总和,包括各种数据库系统、文件系统和应用系统等所能管理的所有数据<sup>[7]</sup>。

数据资产与实物资产具有相似性。数据资产具有价值属性,其价值需要在数据的应用和流通中体现。数据资产按其归属可以分为个人数据资产、企业数据资产和政府数据资产。

定义 2 政府数据资产即现有电子政务系统、规划中的智慧城市工程以及未来建设的政务系统皆属于国有的数据资产<sup>[8-9]</sup>。政府的数据资产与政府的土地资产有相似性,都为国家所有,地方政府可以行使管理权,但可以由政府委托开发商(企业)进行开发。

### 2.2 智慧城市大数据空间

智慧城市的建设除了需要对项目进行规划之外,也需要对数据作总体规划,从顶层设计的高度确定智慧城市数据空间的组成<sup>[10]</sup>。

定义 3 智慧城市大数据空间由智慧城市相关的所有数据集组成,包括所有的数据模型、元数据、原始数据、过程数据等。

定义 4 数据维度即可以将数据空间分解到不同不可交叉的领域并可以进行定性和定量描述的特征集合。

设智慧城市大数据空间为  $\partial$ , 它由多个分属不同领域的空间维度组成。

$$\partial = \{t, s, p, e, o, \partial_1, \partial_2, \dots, \partial_N\} \quad (1)$$

对任意  $\partial_i$  可以下分为不同的子维度。每个数据维度,类似

于一个函数类,其中定义了数据主体特征及其活动。

定义 5 数据抛面即多个大数据空间维度内的对象通过相同含义的属性键值分别相互关联生成的新的数据集。这类类似于某一物体被一个固定的坐标面上切割留下的投影。从数据角度看是某一数据集在多个查询条件限定下的数据实例<sup>[11]</sup>。

对于智慧城市的数据空间,每个空间维度常见的形式组成如表 1 所示。为保证空间的可扩展性,预留其他数据维度接口。如在作相关分析时可以引入气象维度、宏观经济等维度。

表 1 数据空间常见的数据维度

序号	标识	表示	说明
1	$t$	时间维度	发生的时间
2	$s$	空间维度	地理空间位置
3	$p$	人的维度	与人相关的属性及活动集合,下分为市民、职工、操作员等不同的派生维度
4	$e$	物的维度	与物相关的属性及活动集合,下分车辆、设备、建筑等不同的派生维度
5	$o$	组织维度	组织的属性及活动集合,下分企业、政府单位、学校、组等不同的派生维度
6	$c$	安全维度	与安全相关的对象及相关活动,包括常规的密码、手机密保以及指纹、人脸识别等新的安全手段

对于任何  $\forall (\partial_i, \partial_j) \in \partial, \partial_i \cap \partial_j = \emptyset$ , 即任何两个不同的空间维度,不存在数据的交集。

其于数据空间的维度划分,可以形成智慧城市大数据的顶层数据划分及分类模式,梳理围绕不同数据对象的属性及相关活动<sup>[12-13]</sup>。

定义  $\tau(x)$  表示求空间维度的子空间的数量。设时间及空间采用统一的表述模式,从而可以理解  $\tau(t) = 1, \tau(s) = 1$ 。则空间存在的可能的空间交互关系为:

$$L = \tau(p) \times \tau(e) \times \tau(o) \times \dots \times \tau(\partial_j) \quad (2)$$

当然时间的维度并不一定唯一,可以包含农历等不同的历法的维度。对于当前可以定义的空间维度来看,各空间维度的子空间数量是可数的。合理划分空间大小,可以使  $L$  成为一个合理的数值。

### 2.3 业务系统集合

随着时间的积累,在地方政府各委办局中形成了大量的功能及数据相互覆盖的业务系统。定义当前各智慧城市相关的管理部门所有的业务系统集合为  $B$ 。

$$B = \{b_1, b_2, \dots, b_M\} \quad (3)$$

对任意一个业务系统,均体现了几个数据空间维度的相互关联,是其关联集合的向下映射及实例化,也就是一个数据抛面的具体实例化。

各业务系统存在的数据接口数据与业务系统数一致。

$$I = \{I_1, I_2, \dots, I_M\} \quad (4)$$

定义每个业务系数中对应包含的数据集合为:

$$D = \{d_1, d_2, \dots, d_M\} \quad (5)$$

则各数据集在不同的时间、空间等维度的相互关联形成了一个广义上的智慧城市的大数据集,其为智慧城市大数据空间的一个子集。

$$BD = \{d_1 \times d_2 \times \dots \times d_M\} \quad (6)$$

当前可以采集到的大数据集又是  $BD$  的某一个子集。

## 2.4 数据交换目录

用于指导业务系统间数据交换最常用的是数据交换目录,定义与数据接口对应的数据交换目录为:

$$R = \{r_1, r_2, \dots, r_M\} \quad (7)$$

当前数据目录广泛用于电子政务的数据交换。由于缺乏更顶层的抽象,数据目录的扩展及维护缺乏统一的抽象模型,从而阻碍了统一的智慧城市数据库的建设。

综上智慧城市建设必须是在统一的大数据模型下,建设并维护统一的数据空间模型,并建设统一的数据库,相对于不同的业务形成统一的一张表,解决当前数据交换现存的问题。

## 3 智慧城市建设模型

如图 1 所示,智慧城市建设可以概况为“一张网”“一片云”“一张图”“一套库”“一把锁”共 5 个统一的信息化基础设施,从软硬件各个层面为智慧城市建设打下基础。

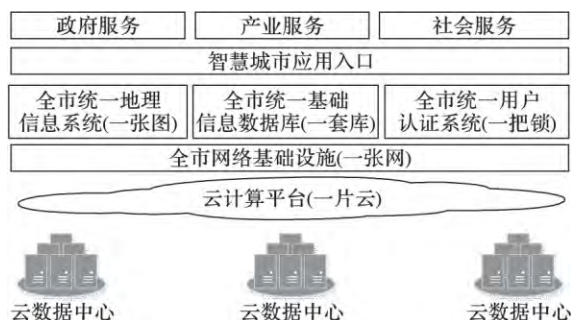


图1 智慧城市建设模型

一片云:建设全市合理分布的云数据中心基础设施,通过云管理平台有效组织,为智慧城市应用提供弹性可扩展的计算、存储和网络资源。

一张网:即建设全市统一规划的高速无线网络基础设施,为各种接入终端、传感器、监控设备提供随时、随地、支持任何装备的全天候的网络接入。

一套库:基于智慧城市统一的大数据模型,建设智慧城市统一的数据库。结合不同的数据空间的抛面生成不同的业务

系统,具有数据抛面交叉的业务系统要共同维护统一的数据表。

一把锁:即建设统一的用户认证及安全管理模型,使用户通过一次登录,可以在安全的前提下使用各类智慧城市应用。

一张图:即选择一个或建设一个全市统一的地理信息空间服务平台,为智慧城市应用提供统一的位置服务。

以上五个部分的内容是智慧城市建设的基本要素。

## 4 建设案例

如图 2 所示为智慧城市大数据开放共享平台建设的一个具体建设案例。智慧城市大数据开放共享平台按照一云、一引擎、四大库、一门户、三大典型应用的思路构建智慧城市大数据开放共享平台<sup>[14]</sup>。

### 4.1 建设框架

大数据开放平享平台的主要内容包括:

一云 基于统一的云管理平台搭建政务服务内网云、外网云、灾备云,建设基于云计算技术的智慧城市大数据开放共享的基础支撑环境;

一引擎 基于用户上下文的授信模型、大数据资源统一注册框架、大数据统一访问接口、大数据统一管理框架、大数据统一业务框架等技术,建设大数据统一驱动及管理引擎,为上层应用提供大数据支撑服务;

四大主题库 基于大数据引擎采集的数据构建智慧城市公开信息数据库、市政地理信息数据库、政企业务信息数据库以及行业领域数据库共四大主题库<sup>[15]</sup>;

一门户 建设一个经过授权及验证的政务服务可信应用门户,发布经审核授权的开发商提供的智慧城市应用,包括移动端和 Web 端;为用户提供安全的应用下载机制,提供用户对应用的评价机制。通过市场机制促进应用的优胜劣汰。

三大应用 通过应用门户提供围绕智慧城市的城市信息公开及定制、城市业务服务、城市公共服务共三大类典型应用示范<sup>[16]</sup>。

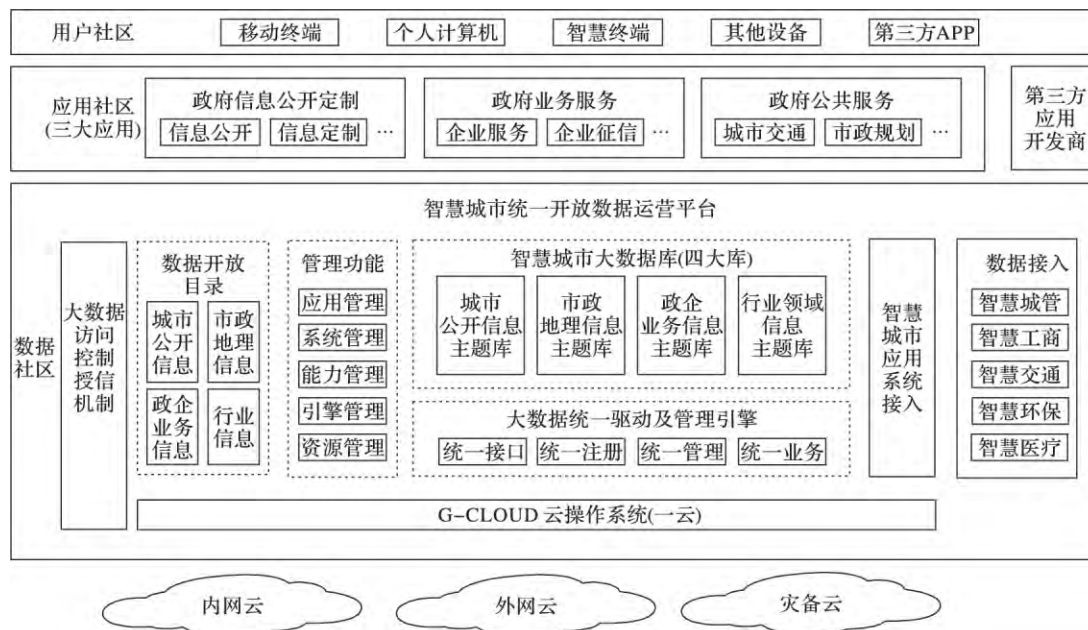


图2 智慧城市大数据开放共享平台建设案例

### 4.2 建设步骤

大数据开放共享平台基于智慧城市云基础设施环境,基

于统一的智慧城市大数据模型实现智慧城市大数据统一管理引擎,建设智慧城市公开信息、政企信息、地理信息、行业信息

等四类大数据主题库。

从安全方面考虑建立智慧城市大数据安全访问授信模型,按授权完成面向不同场景和主体的数据抽取、清洗、统一存储,完成数据脱密、授信及受控访问等安全流程建设;通过大数据统一访问接口,将加工后的数据供政府、企业、个人或第三方应用开发商调用;基于城市用户的庞大用户群,开发面向城市信息公开、政府业务、公共服务三类典型应用,建设一个智慧城市大数据应用发布、数据定制及评价的统一门户,形成数据提供者、应用开发者及用户相互促进的智慧城市大数据共建共享的平台,打造需求、数据、资金合理流向的产业生态<sup>[17]</sup>。

## 5 效益分析

一个城市的全体居民都是智慧城市大数据开放共享平台的潜在用户,因此平台建设居有天然的产业号召力,从而催生数据供应商、资金提供商、应用开发商共同发展的产业生态,并形成产业聚集。

1) 平台建设将进一步推进智慧城市的共建共享。城市居民或企业不再仅仅是平台服务的受众,而是平台建设的共同参与,其价值主要体现在两个层面:一是数据提供者层面,能够为智慧城市大数据开放共享平台提供各种类型数据;二是通过对系统应用的使用和评判,为平台的第三方应用开发商提供商业应用开发环境,形成大数据应用的商业环境。

2) 平台将催生围绕智慧城市大数据的应用创新:一些第三方应用开发商不参与大数据的直接共享,而是基于开放数据开发企业级或用户级应用产品,在为现有数据提供增值的基础上,深度挖掘数据的价值,打造一个围绕平台的产业生态<sup>[18]</sup>。

3) 平台将打造围绕互联网的智慧城市服务“互联网+”已经上升为我国面向全民创新、万众创业的国家战略,也对未来智慧城市建设提供了新的要求和思路。互联网“以用户为中心”的原则,为未来“互联网+智慧城市”勾画了蓝图,即能够以市民、企业等服务对象为中心,建成“一站式”的业务办理、信息查询、信息推送、服务访问等综合型的互联网公共服务平台。

4) 平台将推进城市主管部门的服务创新:智慧城市大数据开放共享平台将进一步发挥互联网、云计算和大数据技术在政府服务职能转换中的作用,推动政府抓住“互联网+”的机遇,形成围绕智慧城市数据运营的公共创新创业平台,建设面向智慧城市的专属开发社区和应用商店,逐步实现“网上政府”一站式服务平台,将为城市的转型升级提供更加有力的支撑<sup>[19]</sup>。

## 6 结语

我国政府各级主管部门都在积极探索围绕智慧城市的建设如何充分发挥市民、企业和其他机构的积极性,形成一个良性互动的城市或城镇化发展模式<sup>[12]</sup>。本文所述的智慧城市的大数据开放平台建设正在推进中,得到了政府和产业界的大力支持。当前智慧城市大数据平台建设缺乏理论的支持和指导。本文描述了智慧城市数据空间的建模模式,从数据维度方面论证了大数据平台的扩展模式,以及现有业务系统和大数据平台的对应关系。本文也描述了一个智慧城市大

数据开放共享平台的建设参考架构,重点围绕智慧城市大数据开放共享平台建设中的难题,给出了一个具体的解决方案。每一个具体的城市的发展阶段不同,会有其文化和产业特色的差别,切入点和道路的选择或将不同。但最终智慧城市大数据平台要建设一个数据创造、数据消费、数据投资的良性的产业循环,形成数据提供者、应用开发者及用户相互促进的依存体系,打造需求、数据、资金合理流向的产业生态<sup>[18-19]</sup>。

参考文献:

- [1] 李德仁,姚远,邵振峰. 智慧城市中的大数据[J]. 武汉大学学报(信息科学版), 2014, 39(6): 631-640.
- [2] 陈铭,王乾晨,张晓海,等. 智慧城市评价指标体系研究——以“智慧南京”建设为例[J]. 城市发展研究, 2011, 18(5): 84-89.
- [3] BATTY M, AXHAUSEN K W, GIANNOTI F, et al. Smart cities of the future [J]. The European Physical Journal Special Topics, 2012, 214(1): 481-518.
- [4] 郭贺铨. 大数据时代的机遇与挑战[J]. 信息安全与通信保密, 2013(3): 9-10.
- [5] KITCHIN R. Big data, new epistemologies and paradigm shifts[J]. Big Data & Society, 2014, 1(1): 1-12.
- [6] 孟小峰,慈祥. 大数据管理: 概念、技术与挑战[J]. 计算机研究与发展, 2013, 50(1): 146-169.
- [7] 徐志伟,谢毅,海沫,等. 人物三元计算中的通用计算账户与个人信息资产代数[J]. 计算机研究与发展, 2013, 50(6): 1135-1146.
- [8] KITCHIN R. Big data and human geography opportunities, challenges and risks[J]. Dialogues in Human Geography, 2013, 3(3): 262-267.
- [9] 于施洋,杨道玲,王璟璇,等. 基于大数据的智慧政府门户: 从理念到实践[J]. 电子政务, 2013(5): 65-74.
- [10] NEIROTTI P, DE MARCO A D, CAGLINO A C, et al. Current trends in smart city initiatives: some stylized facts [J]. Cities, 2014, 38(5): 25-36.
- [11] 甄峰,秦箫. 大数据在智慧城市研究与规划中的应用[J]. 国际城市规划, 2014, 29(6): 44-50.
- [12] 王家耀,邓国臣. 大数据时代的智慧城市[J]. 测绘科学, 2014, 39(5): 3-7.
- [13] 石晓冬. 大数据时代的城乡规划与智慧城市[J]. 城市规划, 2014, 38(3): 48-52.
- [14] FAN W, BIFET A. Mining big data: current status, and forecast to the future[J]. ACM SIGKDD Explorations Newsletter, 2013, 14(2): 1-5.
- [15] 王珊,王会举,覃雄派,等. 架构大数据: 挑战、现状与展望[J]. 计算机学报, 2011, 34(10): 1741-1752.
- [16] 张永民. 智慧城市总体方案[J]. 中国信息界, 2011(3): 12-21.
- [17] SUCIU G, VULPE A, HALUNGA S, et al. Smart cities built on resilient cloud computing and secure Internet of things [C]// Proceedings of the 2013 19th International Conference on Control Systems and Computer Science. Washington, DC: IEEE Computer Society, 2013: 513-518.
- [18] 徐志伟,李国杰. 普惠计算之十二要点[J]. 集成技术, 2012, 1(1): 20-25.
- [19] 许庆瑞,吴志岩,陈力田. 智慧城市的愿景与架构[J]. 管理工程学报, 2012, 26(4): 1-7.