

地名地址信息采集的规则及数据库结构设计

陈照永

江苏煤炭地质物测队

DOI:10.32629/gmsm.v3i2.610

[摘要] 地名地址是基础地理信息。地名地址普查,有利于地方经济社会协调发展,方便人民群众生产生活。地名地址数据缺乏统一的数据分类、地名地址规范化描述等规范指导,近两年“数字城市”、“智慧城市”等建设也多次呼吁制定城市地名地址的数据规范与规程,以统一地名地址数据。笔者结合参与的地名地址普查项目的经验,对地名地址信息采集的规则和数据库结构设计进行简单探讨。

[关键词] 地名; 地址; 兴趣点; 数据库结构设计

引言

地名地址是基础地理信息,地名地址普查与建库是一项公益性、基础性的工作,有利于地方经济社会协调发展,有利于社会交流交往、方便人民群众生产生活,对提高政府管理水平和公共服务能力具有重要的意义。地名地址数据库建设是“城市”建设的基础工作之一,也是“数字城市”公共地理框架数据建设中的重要组成部分。目前由于地名地址数据缺乏统一的数据分类、地名地址规范化描述等规范指导,以及政府与公众对地名地址数据服务的需求已不仅局限于传统的地名地址信息,还对公众服务感兴趣的地名信息提出了更高的要求,导致地名地址数据在提供数据空间匹配、共享交换、查询服务等方面面临很多困难。目前我国各大中城市正在开展地名地址数据标准或规范探索和研究,近两年“数字城市”、“智慧城市”等建设也多次呼吁制定城市地名地址的数据规范与规程,以统一地名地址数据。

1 信息采集的内容与类别

(1) 行政区域地名: 行政区域是指行政区划管辖范围,分区、街道办事处等。(2) 居民地名称: 社区居委会、居住小区、片区等名称。(3) 普查范围内的街、路、巷名称。(4) 各类自然地理实体名称: 山、河流、湖泊等名称。(5) 主要的各专业部门使用的台、站、港、场等名称。(6) 本地的著名的名胜古迹、纪念地、游览地、古遗址和重要的人工建筑物名称。(7) 单位: 党政机关、民间组织、企业事业单位等。

2 信息采集的技术依据

《数字城市地理空间信息公共平台地名/地址分类、描述及编码规则》CH/Z 9010-2011;《地理信息公共服务平台地理实体与地名地址数据规范》CH/Z 9010-2011;《地名分类与类别代码编制规则》GB/T18521-2001;《中华人民共和国行政区划代码》GB/T-2260-2007;《公共地理信息通用符号》GB/T24356-2009;《数字测绘成果质量要求》(GB/T 17941-2008);《基础地理信息公开表示内容的规定》2010年;《地理信息国家标准手册》2004年;《国民经济行业分类》GB/T4754-2011;《国家地名数据库代码编制规则》2010年。

3 地名地址信息采集

3.1 普查对象

(1) 地名信息: 地名采集主要包括自然地名和人文地名,自然地名主要指水系的和陆地地形,水系又包括河流、湖泊、岛屿、泉等,陆地地形包括山、丘陵、山峰、台地等。人文地名主要指行政区域,包括一至四级(乡级)行政区域及其他区域;非行政区域,包括矿区、工业区、开发区、军事区等;居民点,包括一至五级(村级)行政首府、农村居民点、工矿点等;具有地名意义的交通运输设施,包括水上设施、公路设施等;具有地名意义的水利、电力、通信设施;具有地名意义的纪念地、旅游胜地,包括桥梁、

公路环岛、公园、风景名胜等;其他具有地名意义的建筑,包括房屋、亭、广场、城墙等;具有地名意义的单位;具有地名或地址信息的院落。(2) 地址信息: 采集带有街巷、小区(村)、门牌号、楼栋号、单元号、标志物、兴趣点等标准化地址信息。(3) 兴趣点信息: 兴趣点采集主要指党政机关、企事业单位及社会组织、商业服务、交通运输、餐饮住宿、金融保险、旅游服务、房产楼盘、休闲娱乐、医疗卫生、文教科研、生活服务、公共设施共13大类信息的采集。(4) 其他信息: 采集照片、视频、数据来源、采集时间、是否涉密等与地名地址相关的信息。

3.2 外业采集

外业采集信息前的准备:(1) 信息采集以街区为单元,分小组调查。外业采集前将地形图按标准分幅图打印,以街区的范围分给各个小组。(2) 制作外业调查采集表。外业采集表还应填写调查日期、图号、采集员及检查员签字,外业采集时图表上的编号要一致。(3) 数据采集前各小组除了准备地形图和外业调查表外,还应准备笔、文件夹、相机、电池及证明材料等。(4) 外业采集后应和相邻小组做好接边处理工作,做到不重不漏。

3.3 地名信息采集

(1) 自然地名中的水系及陆地地名只采集有名称且具有象征意义的,地理位置标识点一般在地物的标志物或特征点上。(2) 行政区域地名采集点位一般位于当前政区的政府所在地的点位上。(3) 具有地名意义的交通及附属设施需要采集。(4) 具有地名意义的水利、电力、通讯设施需要采集。(5) 具有地名意义的纪念地、旅游胜地需要采集。(6) 具有地名意义的建筑需要采集。(7) 重大危险源、涉密性单位依据国家相关规范要求做好“涉密”标识。

3.4 地址信息采集

(1) 独立门牌均需要采集(如小区单元有独立门牌号,也需要采集)。(2) 小区内单元有独立门牌号,采集的时候标准地址为“省市区+小区名+单元门牌号+楼幢号”。(3) 未公布地址信息(扩建、新建)的单位、商铺,地址信息按照标准地址描述中“政区+路街巷+标志物或兴趣点名”来描述。(4) 如果地址只到街巷,如“XX路XX市场”,也需要作为地址单独采集。(5) 某个区域实体包含多个门牌时,每个门牌都需要采集。(6) 待拆迁地区,如果实地存在地址信息,只需要采集地址信息,对应的单位名称不需要采集,并在备注中说明。(7) 没有地名信息的地址信息也应作为独立的地名地址点存在。

3.5 兴趣点信息采集

(1) 所有涉及独立门面或挂牌的店铺、企事业单位都需要采集,并且兴趣点名均以挂牌名称为主。(2) 兴趣点含有多个门牌时,取中间一个门牌位置作为兴趣点的定位点,同时所有门牌都要作为地址信息采集,地名信息都为该兴趣点的名称。(3) 独栋房屋(写字楼或大厦)内涉及的单位,只根

据大厅挂牌信息采集相关单位, 每个单位对应一个采集点, 其中地址为该单位的实际地址; 楼宇内的市场(或商场), 内部商铺原则上不再细分。如农贸市场、电脑城等。(4) 公交站点以及公交总站应进行采集, 公交站点的命名以实际站点名称命名, 命名规则为: XX公交站, 街道两边公交站点需要同时采集。(5) 地下出入口(地铁等)名称只采集地上的, 地下等内部信息不采集。地铁名称的命名规则为: ***站*号口。(6) 银行与ATM作为兴趣点及地名地址点, 需要分别采集。(7) 采集公用电话亭时, 名称为“电话亭+电话号码”, 地址为“标志物旁电话亭”。(8) 连锁性质的商铺作为兴趣点, 采集的时候需表明连锁号码或分店名。(9) 采集公共厕所时采集其全名。(10) 采集公共自行车站时, 如果该站有名字, 按照名字采集, 如果没有名字, 使用标志物名来表示。(11) 有电动设施或专门有人经营的游乐场必须采集, 按照游乐场实地名称采集, 如果没有名字, 使用“标志物旁游乐场”来表示。(12) 如果当前兴趣点同时属于附录规定的地名信息, 那么需要同时录入地名的分类代码, 如“XX广场”, 即可作为兴趣点, 也可作为地名。(13) 空店面不作为兴趣点采集(如有地址信息, 需要作为地名地址点采集)。(14) 闲置厂房不作为兴趣点采集(如有地址信息, 需要作为地名地址点采集)。

3.6 其他相关信息采集

(1) 标准地址的拆分项采集应由标准地址自动拆分获得, 如行政区划代码、省级行政区名称、市级行政区名称、区级行政区名称、乡镇级行政区名称、社区级行政区名、村名称、街路巷名称、门牌号、楼栋号、单元号、房间号等。(2) 采集地名地址信息时应拍照, 拍照位置是采集对象的正面或全貌, 有门牌号码的还需要将门牌号的细部单独拍照。对于军事单位或其他涉密单位, 不必拍照。(3) 与旅游、风景名胜相关的地名地址采集, 应采集不少于30秒的视频信息, 并做好编号。(4) 如实填写地名地址信息的数据来源、数据来源年代、采集时间等信息, 如果来源于导航数据, 还应填写与导航数据相关导航分类代码、ID等信息。(5) 在采集过程中在其他属性项中无法填写的说明信息应在备注项中说明。

4 GIS 数据库结构设计

4.1 数据分层

数据集(DataSet)	数据类(FeatureClass)	类型标识	图层标识	对应大类
地名地址数据集	地名地址	点	DMDZPT	第1大类
	水系	点	DMDZPT-W	
	地形地貌	点	DMDZPT-T	
	行政区域	点	DMDZPT-A	
	道路交通	点	DMDZPT-R	

4.2 数据库结构设计

(1) 地址信息

序号	英文名称	中文名称	字段类型	字段长度	可否为空	备注
1.	ADDNAME	标准地址	TEXT	100	否	所在地址的结构化描述信息。如果没有地址信息, 此项填“无”;
2.	ADDCODE	标准地址代码	TEXT	17	否	唯一标识一条地名地址的代码; 如果没有地址信息, 此项填“无”;
3.	ADMIN	行政区划代码	TEXT	12	否	
4.	PRONAME	省级行政区名称	TEXT	50	否	包括省、直辖市。
5.	CITNAME	市级行政区名称	TEXT	50	否	包括: 地级市、地区。
6.	COUNAME	区级行政区名称	TEXT	50	可	如“中南区”
7.	TOWNAME	乡镇级行政区名称	TEXT	50	可	包括: 镇、乡、街道。
8.	CMTNAME	社区级行政区名	TEXT	50	可	如“中南社区”。
9.	VILNAME	村名	TEXT	50	可	
10.	RDNAME	街路巷名称	TEXT	50	可	如“上海路”
11.	DOORPLATE	门牌号	TEXT	10	可	
12.	DOORPLATE_PIC	门牌号图片	BLOB		可	门牌号清晰照片, 图片大小不超过1.5M
13.	DRPLATE	楼栋号	TEXT	10	可	
14.	CELID	单元号	TEXT	2	可	如1单元, 应填写“01”(不足2位前面补“0”)
15.	ROOMID	房间号	TEXT	6	可	如“201”
16.	BUILDINGID	建筑物唯一编码	TEXT	20	可	若有多个, 用英文分号隔开。

(2) 地名信息(通用信息)

序号	英文名称	中文名称	字段类型	字段长度	可否为空	备注
17.	NAME	标准地名	TEXT	100	否	
18.	CLASSNAME	地名类别名称	TEXT	50	否	地名所属分类类别名称; 如果没有地名信息, 此项填“无”;
19.	CLASID	地名类别代码	TEXT	4	否	如果没有地名信息, 此项填“无”;
20.	NAMINALP	地名拼音首写字母	TEXT	50	可	
21.	ALSNAME	别名	TEXT	50	可	
22.	SIPNAME	简称	TEXT	50	可	
23.	HISNAME	曾用名(历史名称)	TEXT	100	可	
24.	ENAME	英文名	TEXT	50	可	
25.	ENTIID	地理实体标识码	TEXT	30	可	若有多个, 用英文分号隔开。

(3) 地名信息(兴趣点信息)

序号	英文名称	中文名称	字段类型	字段长度	可否为空	备注
26.	POICODE	兴趣点类别代码	TEXT	6	否	
27.	POINAME	兴趣点类别名称	TEXT	10	否	
28.	ZIP	邮政编码	TEXT	6	可	
29.	TEL	电话	TEXT	50	可	若有多个, 用英文分号隔开
30.	FAX	传真	TEXT	50	可	若有多个, 用英文分号隔开
31.	DESCRIP	概要描述	TEXT	400	可	
32.	HOME	主页	TEXT	50	可	

(4) 数据来源信息

序号	英文名称	中文名称	字段类型	字段长度	可否为空	备注
34.	SOURCE	数据来源	TEXT	50	否	数据来源标志信息, 包括外业采集、导航数据等来源描述。
35.	SOURCEDATE	数据来源年代	DATE	8	可	生产或者更新地名时采用的参考数据现势性(年代)
36.	NVFID	导航数据永久ID	TEXT	20	可	参考导航数据
37.	LEV	排序级别	INT			数字组合。取值区间为: 1-9, 数值越小, 级别越高, 1-3为高等级, 4-6为中等级, 7-9为低等级。

(5) 采集更新信息

序号	英文名称	中文名称	字段类型	字段长度	可否为空	备注
38.	X坐标	X坐标	DOUBLE		可	小数精度0.00
39.	Y坐标	Y坐标	DOUBLE		可	小数精度0.00
40.	COLNAME	采集人员	TEXT	20	可	
41.	PIGUID	采集照片	BLOB		否	图片大小不超过1.5M
42.	AUDICODE	视频编号	TEXT	20	可	若有多个, 用英文分号隔开
43.	COLDATE	采集时间	DATE	8	可	外业调查收集该地址的日期
44.	NAMDATE	命名时间	DATE	8	可	主管部门正式建立标准地址标牌的日期
45.	EDITDATE	更新时间	DATE	8	可	
46.	INVDATA	无效时间	DATE	8	可	
47.	ISSECRET	是否涉密	Boolean			
48.	DOWNLabel	下载标记	Boolean			下载锁定: 1
49.	REMAKE	备注	TEXT	250		

4.3 地名地址描述与编码

(1) 地址描述: 标准地址采用分段组合的方式描述, 由行政区域、基本区域限定物、局部点位置三大类要素构成。〈标准地址〉::=<行政区域名称>[基本区域限定物名称][局部点位置描述]其中: 1)〈行政区域名称〉::=<省级>[地区级][区县级][乡级]。2)〈基本区域限定物名称〉::=<路>|<街>|<巷>|<居民小区>|<村>。3)〈局部点位置描述〉::=<门(楼)址>|<标志物名>|<兴趣点名>。在基本区域限定物地名中, 如遇有使用小区名和街巷名描述均可的情况时, 街巷名优先于小区名; 局部点位置中, 遇有使用标志物名、兴趣点名和门(楼)址描述均可的情况时, 门(楼)址优先标志物名, 标志物名优先兴趣点名。各地址元素的描述规则如下: 政区表示为: ××市××区××街道(乡、镇); 路表示为: ××路(大道); 街巷表示为: ××街(巷、弄); 小区表示为: ××小区(村、苑); 门址表示为: ××号; 楼址表示为: ××幢; 自然村、标志物以及兴趣点名使用权威部门地名地址资

地理信息系统的基本技术与发展动态

夏佼 孔维彬

重庆市勘测院

DOI:10.32629/gmsm.v3i2.581

[摘要] 随着互联网信息科学技术的不断进步与发展,地理信息系统也跟随时代发展技术不断革新。地理信息系统技术已经应用到不同的领域,解决相关的问题,给工作者带来极大的便利,同时也不断完善它。本文主要是介绍地理信息系统的相关概念、地理信息系统的基本技术以及发展现状,让人们充分了解地理信息系统的相关知识,从而促进我国地理信息不断发展。

[关键词] 地理信息系统技术; 地理信息系统; 地理信息系统发展

伴随我国科学技术与经济水平不断发展,我国各个领域逐渐的智能化、科学化。各行各业想要获得优势需要充分利用地理信息系统技术,只要不断的完善与发展地理信息技术就能够为各领域提供技术。目前我国地理信息系统技术还处于初级阶段,还需要专业人士不断的创新与发展,促进地理信息系统技术发展。所以,前提就需要对地理信息系统全面的了解,不断分析,同时对它的发展动态以及趋势也要了解。

1 地理信息系统技术的相关概念

地理信息系统(Geographic Information Systems, 缩写为GIS)是一门技术系统,1967年由罗杰·汤姆林森提出,用于存储,分析和利用加拿大土地统计局收集的数据,并增设了登记分类因素来进行分析。随着计算机技术、经济水平的提高以及空间技术的发展,地理信息系统也不断的发展,逐渐完善。地理信息系统是用于输入、分析、查询、存储以及显示地理数据的计算机系统,同时也是一门综合性的学科,结合了地理学和地图学,目前已经应用在各行业各样,在解决问题的同时也逐渐完善。它主要包括人员、数据、硬件、软件、过程五个部分。其中,人员是地理信息系统最重要的组成部分。信息技术的不断进步,世界各国也制造出相应的地理信息系统软件,按照地域范围可以将其划分为全球性、区域性、局部性三方面。除此之外,我们还要清楚一点,那就是地理信息系统中“地理”不是我们所理解的狭义的地理学而是广义上的地理,主要是指地地理坐标参照系统中的空间数据、属性数据以及在此基础上得到的相关数据^[1]。因此,地理信息系统人们利用计算机硬件与软件技术对数据库进行分析、科学操作管理,能够将庞大的数据信息经过一系列的操作,比如:空间函数等技术方法,能够呈现出一个直观的数据现象,对整体以及各个部分进行总结分析和预测,从而产生有利的信息数据,促进各领域的发展,使得地理信息系统朝着

料中已有的名称。(2)地址编码:编码规则如下图所示:1)行政区域代码:参照最新的《南京行政区划代码》,参照附录A中的编码表。其中县级及县级以上行政区代码为1-6位,乡镇街道行政区代码7-9位,编码规则参照GB/T 10114-2003《县级以下行政区划代码编制规则》;社区行政区代码为第10-11位,可以按照乡镇街道下的社区顺序编制。2)基本区域限定物代码:四位全部为“0”。3)局部点代码:使用流水号表示。(3)地名类别代码:地名类别代码分为门类、大类、中类、小类四层,分别用1位阿拉伯数字表示,如果中类不再细分,则它们的末位用“0”补齐。大类、中类、小类3层中的“其他”类别用“9”表示。(4)兴趣点类别代码:编码规则如下图所示:地名类别代码分为大类、小类、细分类三层,分别用2位阿拉伯数字表示,大类和小类都从“1”顺序编号,不足两位数的前面用“0”补齐,其他的小类,统一用“99”表示,细分类主要为将来扩展细化,目前统一以“00”表示。

5 结语

智能化、科学化、集成化、标准化等方向发展。

2 地理信息系统的基本技术

2.1 数据输入与编辑技术

输入主要是指地理信息系统空间信息数据库的输入,预处理的过程。空间数据只有加工成合适的形式才能被地理信息系统所利用,而我们把空间数据输入计算机的过程被称为数字化。目前我们最常使用的方式有数字化系统和扫描式。数字化的系统与其他相比较而言效率比较低,成本高,而扫描式的输入方式不具有实时性,也就是没有不可能及时转换为所需要的数据信息。目前我们根据数据信息的来源以及处理的方式,使用了较为普通的矢量方法完成交互式的识别。

地理信息系统技术的数据编辑主要表现在属性编辑与图形编辑。属性编辑需要与相应的数据库管理有机结合共同完成的。而图形编辑具有拓扑关系的建立、图形编辑、图形整理、图幅拼接、图形变换、投影变化、误差修正等功能,为处理数据提供了便捷途径。

2.2 数据存储与管理技术

数据存储主要是将数据以某种方式存储在计算机内容或者外部介质上,为数据库信息的存放提供位置,方便对数据进行操作。数据在计算机存储也有特定的方式,主要是根据文件结构确定的,同时也会确立与之相对应的逻辑顺序。地理信息系统不仅能够对普通的文件结构构建数据,更重要的是能够复杂且大型的数据库系统进行严格有效的管理。比如,在关系数据库系统中,我们可以通过相关表格中的主码与外码连接,提高地理信息系统的扩展性以及灵活性。

地理信息系统数据的管理主要是包括数据查询与检索、分析与数据显示与输出。数据查询与检索主要从数据库中或者空间信息中找到用户所需

统一地名地址数据规范以及生产采集原则与要求,利于建设标准统一的地名地址数据库,以便于今后提供数据查询、定位、统计以及专题信息空间匹配服务,也有利于各类空间信息平台共享交换。

[参考文献]

- [1] GB/T 2260-2007, 中华人民共和国行政区划代码[S].北京:中国标准出版社,2008.
- [2] GB/T 18521-2001, 地名分类与类别代码编制规则[S].北京:中国标准出版社,2002.
- [3] 王平,薄正权.地名地址数据采集方法与实践[J].城市勘测,2013(2):54-57.
- [4] 马学峰.湛江市地名地址数据库设计与实现[J].测绘通报,2014(S2):288-291.
- [5] CH/Z 9002-2007, 数字城市地理空间信息公共平台地名/地址分类、描述及编码规则[S].北京:测绘出版社,2008.