

地理实体分类与编码体系的构建

阎国年^{1,2}, 俞肇元^{1,2}, 周良辰^{1,2}, 兰馨^{1,2}

(1.虚拟地理环境教育部重点实验室(南京师范大学),江苏 南京 210023;

2.江苏省地理信息产业研究院,江苏 南京 210023)

摘要 地理实体是政府管理的基础单元,是各类信息聚合的凝聚核。地理实体的构建是“数字城市”和“智慧城市”建设的重要内容。本文针对地理实体分类体系的多样性、粒度划分的层次性、描述表达的完备性,以及实体编码的唯一性等关键问题,提出了面向“智慧城市”建设的地理实体科学分类体系,应用需求驱动的地理实体颗粒度分割体系,地理视角的地理实体描述体系,以及顾及地理实体空间分布、时间过程的唯一身份编码体系构建的设想与思路,为地理实体建设工程的顺利实施提供理论依据和方法途径。

关键词 地理实体; 分类体系; 编码体系; 颗粒度

中图分类号: P208.2

文献标识码: A

文章编号: 1672-4097(2019)01-0001-06

0 引言

地理实体是人类为了描述和表达地理世界中具有特定空间范围、形态、过程、关系,以及相关属性地理现象的实体化抽象。同时,地理实体也是政府管理的基础单元和基本内容,以及人居环境的载体和人类活动的基本对象。得益于地理实体能够有效承载着各类地理信息,作为地理信息整合和共享交换的“凝聚核”,其在城市空间规划、建设与管理,数字城市、智慧城市建设中具有关键作用。正因如此,地理实体的相关工作应成为当今地理信息数据资源建设、地理信息系统建设、地理信息应用的重点研究内容。然而,由于地理实体存在多样性和复杂性等特点,目前对地理实体的分类体系、颗粒度分割、描述与表达体系的研究还相对薄弱。在地理实体的唯一编码(地理实体身份证)实现上面临着较大的困难。这直接导致了与地理实体相关信息聚合和共享交换的困难,以及有限且不完备的基础地理实体信息对“数字城市”和“智慧城市”建设的支撑仍相对薄弱。

1 地理实体分类体系多样性

地理实体(GeoEntity)是独立存在,并可唯一性标识的自然或人工地物目标和对象^[1],可以是客观物理存在的,如道路、水体、建构物等,也可以是人类抽象虚拟存在的,如经纬线、等高线、等值面、行政区、管理区等^[2]。

不同的视角、学科、行业及应用需求都可以形成对地理世界中的不同地理实体分类。其中,从地

理实体分类的维度上看,可以存在从学科视角、领域视角、主导因子视角、应用领域视角和时间演化视角等多个不同的分类角度。以地质学相关的地理实体为例,其既可以从岩石学、地层古生物学、构造地质学、地球物理学、地球化学等学科视角和水文地质、工程地质、环境地质、灾害地质等领域视角进行分类,又可以从大地构造、区域构造等不同空间尺度和从板块构造、地质力学、多旋回构造运动、地洼(活化)构造等不同机理以及不同的地质资源类型视角等进行分类。

对于同一学科的地理实体,其根据观察视角和主控因子的不同往往呈现不同的划分依据。如综合自然地理学中,黄秉维运用地带性规律,将我国划分出6个热量带1个大区18个地区和亚地区,28个地带,88个自然省^[3]。任美铎等依据自然差异的主要矛盾以及利用、改造自然的不同方向,将我国划分为8个自然区,23个自然地区和65个自然省。1978年修改为8个区,28个亚区,58个小区。1988年进一步修改为8个自然区,30个自然亚区,71个自然小区^[4]。在人文地理学中,方创琳等以自然、经济、人口、文化、民族、农业、交通、城镇化、聚落景观和行政区划10大要素为基础基础,以主导因子,自然、行政和文化景观单元为基础,将中国人文地理划分为8个人文地理大区 and 66个人文地理区(不含台湾省)^[5]。

类似的,在同一学科中,根据地理实体的属性差异也可以区分出不同颗粒度的地理实体。如在城市路灯管理中,按路灯高度分为高杆灯、中杆灯、道路灯、庭院灯、草坪灯、地埋灯;按路灯杆材质分

项目来源:国家自然科学基金重点项目(41231173)

第一作者简介:阎国年,教授,虚拟地理环境教育部重点实验室主任,主要研究方向为地理信息科学。

为热镀锌铁质路灯,热镀锌钢质路灯与不锈钢路灯;按路灯光源分为钠灯路灯,LED路灯,节能路灯、新型索明氙气路灯;按造型分为中华灯、仿古灯、景观灯、单臂路灯,双臂路灯;按供电方式分为市电路灯、太阳能路灯、风光互补路灯等。

需要明确的是,地理实体的划分并非一成不变,其往往会随着时间的不同而发生变化。如在土地利用现状的分类方面,2007年土地利用现状分类国家标准(GB/T 21010—2007)主要依据土地利用现状、用途、经营特点与覆盖特征,将土地利用现状分类分为8个一级类,40个二级类;2017年土地利用现状分类国家标准(GB/T 21010—2017)则将土地现状分为12个一级类,73个二级类;2018年国土三调工作又将土地利用分为13个一级类,55个二级类。

上述分析显示,由于地理实体在学科、领域、颗粒度、应用范围以及时间上的多样性和差异性,使得实际使用的地理实体的分类体系多种多样,不同学科之间、行业之间、不同应用需求之间,乃至科学家群体之间对地理实体的分类不一致。虽然经过学科的争鸣、行业的规范,在一段时间内形成了一些标准规范,但是地理实体分类体系上的差异性则会始终存在。其核心原因在于现有的地理实体分类都是自下而上的归纳方法构建的,不同领域之间未能进行高效的贯通。

地理学是地学研究的集大成者,以宏观、系统、综合的方法论见长。从地理学学科体系的内在规律性出发对地理实体进行顶层设计,可以自上而下的形成以地理学学科体系为指导的地理实体分类体系(表1)。以地理学学科分类的高度来统筹地理实体的分类,这一分类范式使得实体间的关系明晰、逻辑性强。这种分类还可以纲举目张,能有效地防止以不同应用需求为驱动的地理实体分类而出现挂一漏万的现象,并导致出现重大偏差。

地名分类与类别代码编制规则(GB/T 18521—2001)将地理实体分为自然地理实体和人文地理实体,这一分类基本符合地理学对地理实体的科学划分,但是其在对自然地理实体的分类时只是考虑了地貌和有限的土地利用现状类型,与自然地理学对地理实体的划分差距甚大。而人文地理实体主要考虑了“人造地理实体”,仍有大量的人文地理实体没有考虑。同时,伴随着信息化时代的快速到来,以物理空间、人文社会空间和信息空间的深度融合,人一机一物的协同互联使得信息空间中的虚拟对象也需要进入地理实体的管理体系^[6-7]。本文提出的地理实体分类体系对自然地理、人文地理和信息地理学三个学科体系进行了系统梳理与重构,例

如在自然地理学方面增加了基于地貌学的自然地理实体(包括地形地貌、构造地貌、风化剥蚀地貌、重力地貌、流水地貌、冰川地貌、风力地貌(风蚀地貌、风沙地貌)、冻融地貌、喀斯特地貌、黄土地貌、丹霞地貌等)、基于自然资源的自然地理实体(包括土地、地质矿产、地形地貌、水、土壤、气候、生物等)和基于自然灾害的自然地理实体(包括地质灾害、气象灾害、洪涝灾害、土壤退化、病虫害等)。在人文地理学方面,增加了基于经济地理学的人文地理实体(包括农业、工业、交通运输、商业、金融、贸易等)和基于政治地理学的人文地理实体等(包括军

表1 地理学学科分类与地理实体划分

自然地理学	部门自然地理学	地质学 地貌学 水文地理学 土壤地理学 气候气象学 生物地理学
	综合自然地理学	自然环境 自然资源 自然事件 自然灾害 历史自然地理学 自然地理综合
	区域自然地理学	世界自然地理 海洋地理 大洲自然地理 国家自然地理 区域自然地理
	部门人文地理学	经济地理学 社会地理学 文化地理学 政治地理学 旅游地理学 人口地理学
人文地理学	综合人文地理学	人文资源 人文环境 人文事件 历史人文地理 土地资源 人文地理综合
	区域人文地理学	世界地理 区域地理 城市地理 乡村地理
	部门信息地理学	时间地理学 空间地理学 对象地理学 现象地理学 事件地理学 场景地理学
信息地理学	综合信息地理学	信息基础设施 信息地理综合
	区域信息地理学	海洋信息地理学 陆地信息地理学 天空信息地理学 网络空间信息地理学 虚拟空间信息地理学

事、地缘政治、行政区划等)。表1也给出了信息地理学的分类体系,并从信息空间中对象的类型、性质、属性及其相互作用关系等视角对信息地理学中的地理实体对象进行了系统划分。

2 地理实体粒度划分层次性

地理实体的分类具有层次等级的划分,不同层次等级的分类则代表了地理实体颗粒的大小。如《中华人民共和国宪法明确规定了目前我国的行政区划,全国分为省、自治区、直辖市、特别行政区一级行政区,地区、地级市、自治州等二级行政区,区、县(县级市)、自治县等三级行政区,街道、乡、民族乡、镇等四级行政区,乡镇是我国最基层的行政单位。然而,为了便于政府管理和社会治理,在行政区划的基础上,地名分类与类别代码规则(GB/T18521-2001)又进一步划分出区域性群众自治组织辖区,如村民委员会辖区、居民委员会辖区。为了进行有效的城市管理,城市市政综合监管信息系统管理部件和事件分类、编码及数据要求(CJ/H214-2007)规定在街道的基础上进一步划分单元网格。为了满足城市不动产管理与人口管理要求,在街道的基础上,又进一步划分出街坊、宗地、幢、层、户、室、房间等地理实体。

同类地理实体具有层次等级鲜明的特征,而不同类型的地理实体之间在同一等级或层次中往往存在交叉切割的问题,造成地理实体粒度产生“变异”。例如,当以河流的中心线作为行政区、管理区的边界,这时就要求对河流地理实体的分割,这就不像以单线河、双线河表达那样简单,而是要以河流的中泓线进行河流实体的进一步分割。当道路中心线作为行政区、管理区的边界时,就要求利用道路地理实体的中心线、道路中间的隔离带、护栏等进一步将道路分割为左行道、右行道。当进行道路交通管理时,又要将左行、右行道路实体进一步划分直行、左转、右转不同的车道实体。

上述交叉切割存在的问题源于地理实体的粒度划分层次性、颗粒度的大小取决于人类对地理世界的认知程度、地理实体的管理需求。目前,我国各行各业都在按照自身的管理需求,进行地理实体的普查建库。这就要求对地理实体颗粒度的划分时,需要在对地理实体以科学分类的基础上,统筹现有同类地理实体不同的权威管理部门制定的标准规范,形成不同尺度的基础地理实体的最小颗粒。在此基础上,再面向不同行业、不同领域、不同管理层级的应用层次进行进一步的颗粒划分,确定各自的最小管理单元、最小管理颗粒,形成自身面

向管理的“地理实体最小颗粒”。

“地理最小颗粒”是相对于不同的行业、领域、管理层级管理需求的。地理实体粒度如果划分过粗,则不能满足精细尺度的部门要求,而粒度划分过细,工作量又太大,造成资源浪费。当前,我们可以通过对不同级别政府管理部门对地理实体管理的职能分析,得出不同地理实体颗粒可共享的程度,并依据共享程度确定哪些地理实体类型、哪些地理实体颗粒数据的生产秩序。值得指出的是,并非所有的行业、领域、管理层级都需要完整构建自己地理实体颗粒分类体系,很多地理实体是多个部门共有的,只有少量的地理实体是少数部门共有或只有某些部门独有的。因而,对于地理实体颗粒分类体系的构建实际上主要解决的是不同行业、领域和管理层级等多个部门间的地理实体如何实现共享的问题。

地理实体颗粒的共享需要解决的关键问题是基础地理实体与专题地理实体同步更新的机制问题。由于社会对地理实体数据共享需求迫切,以及基础地理实体数据生产效率的提高,基础地理实体数据更新频率越来越高,从几年、一年、半年到三个月。基础地理实体数据更新需要采用增量更新方法,而专题地理实体数据则在基础地理实体增量变化的基础上进行更新,从而有效减少专题地理实体数据生产更新的工作量,防止因为基础地理实体数据整体更新造成专题地理实体数据来不及更新,而造成的不愿意基于基础地理实体数据基础上进行更新,避免基础地理实体数据又一次沦为“背景图”。

3 地理实体表述的完备性

地理实体自被人类认识、表达开始,就以一定的形态存在,由其几何形态、语义特征和属性信息加以描述的^[8],并由表示该实体与环境中的其他实体联系的关系信息共同构成^[9]。过去,我们将地理信息内涵界定为空间和属性信息。其中,地理实体空间信息的描述主要包括位置信息和几何信息。位置信息通常采用绝对空间位置和相对空间位置来描述。对于绝对位置的描述,一般情况下以平面控制网和高程控制网为基础,将地球上的自然表面转移到椭球面上,从而确定地理实体的地理空间位置(经纬度坐标),并可通过投影换算成平面直角坐标(公里坐标);对于相对位置的描述则是将不同地理实体相互之间的距离关系、方位关系、拓扑关系进行相对表达。如使用线性参照系进行高速公路里程表达等。对地理实体的几何信息表达普遍采

用点、线、面、体、像素、体素等形式,用坐标和坐标串加以显式或隐式表达。由于对地理实体的图形表达受制于制图比例尺,在大比例尺地图上丰富多样的地理实体在小比例尺地图上不同形态、不同大小、不同分类内容加以抽象概括为地理实体。

在地理实体的属性描述方面,现有地理实体描述主要是基于地图学视角将地理实体的部分语义信息,包括名称、分类、编码;几何信息,包括位置、长度、面积、体积;过程信息,包括时刻点的状态、变动趋势;部分的属性信息,包括物理、化学、生物、人文、社会、经济等均作为属性加以描述的。

然而对地理实体以空间位置、几何形态和相关属性只是对地理实体几个侧面的表达,单纯的属性字段并不能准确的反映地理实体的真实状态,难以形成对地理实体语义特征、演化过程、形成机制的系统刻画与整合描述,从而导致地理实体在采集、管理、共享和应用等方面的一系列问题。在人人都是地理实体信息采集员、加工者和使用者,以及机器也是信息采集员和使用者的今天,必须对地理实体信息进行更为系统的划分。因此,作者结合前期在地理场景学等方面的工作^[10,11],提出地理实体的完备表述应该从语义描述、空间位置、几何形态、演化过程、相互关系和属性特征等方面进行(表2)。

表2 地理实体的描述体系

语义描述	地理实体名称解释、分类体系位置、原理图
空间定位	绝对位置、相对位置、地名地址
几何形态	点、线、面、体、像素、体素
演化过程	时刻点及其状态,时间片段与生命周期中的行为、事件
要素关系	空间关系、时间关系、地理要素间相互关系
属性特征	物理、化学、生物、人文、社会、经济相关属性

与地图学视角的地理实体的描述内容相比,地理学视角对地理实体的描述强化了语义描述。过去GIS缺少语义,所以有一段时间学界强调要发展语义GIS。学者们一致认为,在地理实体的语义描述内容中不仅要给出其名称,而且要对其名称进行解释。同时,要给出地理实体对应的分类体系,以及该实体在分类体系中的位置,即语义位置。其中,最为重要的是对地理实体表达的原理图构建,这是该实体结构、功能、规律的最高层次的图形化抽象。由此实现地理实体语义的可读性、分类的可定位性、结构的可抽象性、功能的可视性,以及规律的认知性。

传统地理实体缺少时间过程的动态性、变化过程的多态性、以及地理事件驱动性描述,地理学视

角对地理实体的描述增加了地理实体演化过程的表达。首先是对时间的表达,采用了时刻点、时间片段和生命周期全过程对时间进行刻画,同时对地理实体在某一时刻点的状态,在某一时间片段、全生命周期内的行为描述,特别对地理实体在演化过程中所发生的事件进行表达。例如,根据地理实体的生命周期,我们可以将地理实体分为现状地理实体、规划地理实体、建设地理实体和管理地理实体。显而易见,现状地理实体已经存在的地理实体,规划地理实体为通过各类规划得到的地理实体,建设地理实体为施工建设过程中所涉及到的地理实体,而管理地理实体则为不同政府部门所管理的地理实体。

在要素关系方面,不仅刻画地理实体的语义空间关系,如距离关系、方位关系、拓扑关系等,还要描述地理实体演化过程的时间关系,如顺序的、倒序的、插补的关系,更重要的是表达地理实体内部、实体之间的相互联系与相互作用,如地质如何影响地形地貌、地形地貌如何影响气候条件、气候如何影响水文、水文如何影响土壤、水土如何影响生物等生物地球化过程,以及人类活动如何地球系统相互作用及其作用机制。

在属性信息方面,除了地理实体几何属性描述外,更多强调了其在物理学、化学、生物学方面,以及人文、社会、经济方面的信息表达,辅以刻画地理实体的相互作用机制,包括物理、化学、生物学作用,以及人文、社会、经济作用所表现的特性。

4 地理实体编码的唯一性

地理实体编码是为识别地理实体而设置的编码体系,它将全部地理实体按照分类标准、层次等级颗粒进行编码。目前的相关地理信息编码、地理实体编码有很多国家标准、行业标准、地方标准,这些标准规范中给出的编码、代码主要为类型码。在基础地理信息系统、各类专题地理信息系统建设时,对基础地理实体、专题地理实体都需要进行唯一编码,这种编码常常是在类型码之后加上几位计算机自动生成的顺序码,这种编码会随着地理实体的增加、删除、修改而发生变化,因而不能作为每个地理实体的唯一代码。有些专题性的地理信息系统对地理实体的编码仅在某一区域上、某一时间段内做到了唯一代码,如县级以上行政区划编码,某一个城市部件的编码,某一条输电线路电力设施的编码等。

目前,地理实体的编码少有像居民身份证那样,由地址码、出生日期码、顺序码和校验码构成,

这就使得地理实体编码唯一性难以实现。没有唯一的身份编码,导致了实体信息共享,数据交换,数据关联,查询检索都会出现问题。参考居民身份证对信息编码的结构,本文提出了地理实体身份编码的设计方案(图 1),该编码方案充分考虑了地理实体所在的空间区域、类型等级、图形包络、产生时间等,由大洲或国际公有领域 1 位代码,国家和地区名称 3 位代码,行政区划 6 位代码,县级以下行政区划

3 位代码,地理实体类型 8 位代码,空间网格 8 位代码,生命周期 8 位代码,顺序码 4 位,校验码 1 位构成。上述地理实体唯一编码方案可以很好地实现地理实体的唯一身份识别,并能从编码上直接获得地理实体的语义信息,无论是对于人—人、人—机、机—机之间的信息处理与交换都具有明显的促进作用,有助于应对人机物万物互联条件下的地理实体的采集、识别、处理、分析与应用。

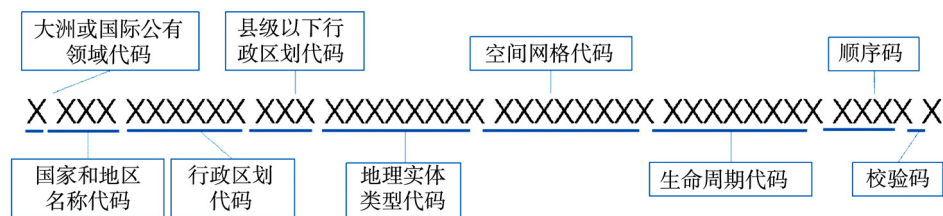


图 1 地理实体唯一编码方案

5 结 语

地理信息已经从对几何对象的描述扩展和提升到地理实体表达,信息聚合已经从基于地理空间位置的聚合发展到基于地理实体的聚合。地理实体作为构建地理场景的重要组成部分和政府管理的对象,在“数字城市”地理空间框架建设过程中得以重视,并将在“智慧城市”建设过程得到进一步强化。对地理实体分类体系的构建,必须认识到地理实体所具有的客观性,更要重视其主观性,取决于不同的视角、不同的学科、不同的行业、不同的应用需求。为了尽可能解决地理实体多样性存在的问题,需要以地理学理论为指导,在地理学学科分类的基础上进行宏观层次上的分类,从而有效解决分类的科学性问题。对地理实体粒度的划分与最小颗粒度的确定,需要认识到粒度划分、粒度大小取决于人类对地理世界的认知程度,特别是对地理实体的管理需求。需要统筹现有同类地理实体不同的权威管理部门制定的标准规范,形成不同尺度的基础地理实体、面向行业的基础地理实体、行业专题地理实体分级体系,以及不同行业的最小地理实体颗粒,实现地理实体不同层次和颗粒度的共享交换。对地理实体的描述与表达,需要在“空间+属性”描述的基础上,以地理学视角,发展地理实体描述的 6 个维度,即语义、位置、几何、过程、关系和属性,实现地理实体的全空间^[12,13]、全信息^[11,14,15]表述,形成新型的 GIS 数据模型,有效解决大数据在空间位置与地理实体的聚合问题。对地理实体的唯一身份编码,需要注意利用地理实体的空间分布特征和演化过程规律,选择和使用合适的空间网格,形成包括对地理实体空间区域、类型等级、图形

包络、产生时间编码,设计唯编码,给出身份信息,这样才能实现地理实体的信息共享,数据交换,查询检索,有效支撑“智慧城市”建设。

参考文献

- [1] Yuan M. Representing complex geographic phenomena in GIS [J]. Cartography and Geographic Information Science, 2001 Jan 1;28(2):83-96.
- [2] 李景文,刘军锋,董星星.基于实体的地理空间数据模型描述与表达[J].测绘与空间地理信息,2008,31(6):1-3.
- [3] 黄秉维.中国综合自然区划草案[J].科学通报,1959,18:594-602.
- [4] 任美镔,杨纫章,包浩生.中国自然区划纲要[J].北京:商务印书馆,1979.
- [5] 方创琳,刘海猛,罗奎,等.中国人文地理综合区划[J].地理学报,2017,72(2):179-196.
- [6] Alex de Freitas C. Changing spaces: Locating public space at the intersection of the physical and digital [J]. Geography Compass, 2010, 4(6): 630-643.
- [7] Batty, M. Digital twins. Environment and Planning B: Urban Analytics and City Science, 2018,45(5): 817-820.
- [8] 周成虎,裴韬.地理信息系统空间分析原理[M].北京:科学出版社,2011:1-3.
- [9] 邬伦,王晓明,高勇,等.基于地理认知的 GIS 数据元模型研究[J].遥感学报,2005(5):583-588.
- [10] 阎国年,俞肇元,袁林旺,等.地图学的未来是场景学吗?[J].地球信息科学学报,2018,20(1):1-6.
- [11] 阎国年,袁林旺,俞肇元.地理学视角下测绘地理信息再透视[J].测绘学报,2017,46(10):1549-1556.
- [12] 周成虎.全空间地理信息系统展望[J].地理科学进展,2015,34(2):129-131.

- [13] 华一新.全空间信息系统的核心问题和关键技术[J].测绘科学技术学报,2016,33(4):331—335.
- [14] 周成虎.全息地图时代已经来临——地图功能的历史演变[J].测绘科学,39(7):3—8.
- [15] 周成虎,朱欣焰,王蒙,等.全息位置地图研究[J].地理科学进展,30(11):1331—1335.

Construction of Geographical Entity Classification and Coding System

LÜ Guo-nian^{1,2}, YU Zhao-yuan^{1,2}, ZHOU Liang-chen^{1,2}, LAN Xin^{1,2}

(1. Key Lab of Virtual Geographic Environment (Nanjing Normal University) MOE, Nanjing Jiangsu 210023, China; 2. Jiangsu Institute of Geographic Information Industry, Nanjing Jiangsu 210023, China)

Abstract Geographical entities are the basic unit of government management and are the condensed cores of various types of information aggregation. The construction of geographic entities is an important part of the construction of “digital cities” and “smart cities”. Based on the diversity of geographic entity classification system, the hierarchy of granularity, the completeness of description and the uniqueness of entity coding, this paper proposes a scientific classification system of geographic entities for the construction of “smart city”. The geographic entity granularity segmentation system, the geographic entity description system of geographic perspective, and the assumptions and ideas of the unique identity coding system construction considering the spatial distribution and time process of geographic entities provide theoretical basis and methodological approach for the smooth implementation of geographic entity construction projects.

Key words geographic entity; classification system; coding system; granularity send feedback