

国土资源信息标准化与信息共享

Information Standardization of Territorial Resources and Information Sharing

王正兴

(中国科学院自然资源综合考察委员会 北京 100101)

国土资源信息特征

一般而言,国土资源信息具有以下六个特征:

1. 空间上的地域性 由于国土资源本身的地带性和非地带性特点,用于描述它们的信息也有地域性特征。这就需要进行各种综合的或专项的区划、规划等。在以计算机为基础的信息系统中,只需给数据项增加相应“分区码”即可描述这种地域性。但在实际工作中,由于各用户对这种地域性认识不一致,致使冲突在数据输入计算机之前即已存在。

2. 时间上的多变性 由于国土资源许多要素不断演变,客观上要求国土资源信息具有一定“时效性”,要适时进行“更新”。利用同一点不同时间的信息可分析国土资源的动态特征。

3. 边界模糊性和分类不确定性 虽然存在“地域性”,但国土资源的地域是边界模糊的,如黄土高原与其北部风沙区并不存在明显界限。同理,地面坡度是客观存在的,它对资源利用的制约也是有规可循的。但是,如何定量地描述这种模糊性和不确定性,不同学者意见或许有分异,但这种分异也未必总能分出优劣高下。在学术自由的环境下,如果能逐渐形成权威当然最佳,否则,“缺乏共同语言”势必影响学术交流和信息共享。

4. 应用上的综合性 利用信息决策时,多是要综合考虑各方不同的信息。如综合农业开发项目需要使用土地、劳力、市场、环境等方面的信息,如果各种信息不规范,决策质量将受到影响。

5. 有价值性和公益性 国土资源信息是国家投资获取的,它当然是有价的,它的使用价值是促进国家资源的保护和持续利用,但其价值的实现一般不是通过买卖关系实现的,因为国土资源信息的主要用户还是国家自己,所以,国土资源信息一般是“公共资源”,具有“公益性”。

6. 可共享性 从是否可共享看,资源可分为“排它性”和“可共享性”。一般资源,一人消费,另一人便不能消费;即使是可再生资源,这一代消费后,必须等若干年才能再次消费。信息资源也有这类特

征:排它性信息如商业秘密等,只有独占才能获暴利,而可共享的信息如国土资源,信息使用的人越多,这一资源的投入产出比越高,国家财富增进得越多。

从以上国土信息的六个特征可看出,要充分发挥信息本身潜能,必须首先消除信息本身可能存在的天然弱点,提高信息质量,以信息标准化为基础促进信息共享。以下分别从信息本身和信息用户两方面讨论目前信息共享中的问题。

国土资源信息质量要素

Stan Aronoff(1989)从三方面讨论了信息质量:微观尺度、宏观尺度和使用难易程度。

1. 微观要素:包括地理位置准确性、属性精确性、逻辑一致性、分辨率等 空间位置准确性指选择合适的坐标系统,并具备一定的精度。属性精确性指分类多少,术语及分类的标准化等。逻辑一致性指空间同一点在某一时刻属性的一致性:如果已评为宜林地,就不应再统计为宜牧地,以免重复计算。分辨率指最小分类单位,如图像像元大小、最小制图单位、最小分类等。高质量信息允许用户以更灵活的方式对信息进行“二次开发”,但高质量往往需要较高投入。

2. 宏观要素:包括完整性、时效性和背景

(1)信息完整性 在空间上表现为对感兴趣地区的完整覆盖;在分类上表现为对感兴趣对象的穷尽分类;在精度上表现为检验数据的完整性。这里只对空间完整性略作讨论。由于资源科学尚处于初创期,还未形成公认的区划方案,所以,各有关专业在进行新课题设置时,一般不是采用已有的区划方案,而是自己重新开始分区。这便形成了“一事一分区”现象。由于前面讨论的“模糊性”,这样的分区往往与前人的分区是不兼容的,而根据“一事一分区”现象获取的信息对后来者——如果恰好也奉行“一事一分区”政策——便很少有参考价值。于是,在一相当长时间内,对同一地理单元(至少核心区不变,

如黄土高原),尽管各方已做了大量工作,积累了大量信息,但当综合分析这些信息时,便会遇到许多困难,如信息完整问题。有时,信息是以较小的地理单元(如县)发布的,通过缩小感兴趣地区也可作些综合分析。但更常见的信息是以“一事一分区”的系统发布的,这样的信息格式限制了其它用户的“二次开发”,限制了共享。

(2)时效性 信息价值随时间而变,在我们这样一个国土辽阔、技术相对落后、信息相对贫乏的国家,即使采用遥感等先进技术,要想对基本国土资源信息在县级水平上更新一次,也需要几年时间。再加上许多难以预料的实际问题,从解译信息源到正式发布,某些信息可能已无多大实用价值。这方面的教训很多。

(3)信息背景 即“有关信息的信息”。任何信息都首先来自某种信息源,经过某种方式加工后到达用户手中。理想的信息应包括背景信息,这可帮助用户更深入地了解信息质量,以便更好地使用已有信息;但一般用户往往很少获得信息背景。

3. 使用难易程度:包括信息可获取性和费用
如果信息的其它方面都是完美无缺的,但因某种原因使用户在获取和使用困难重重,这样的信息也存在质量问题。这一问题在下文中进一步讨论。

用户遇到的信息共享问题

除以上信息本身可能存在的问题外,从用户使用信息的角度也可看到信息管理中不同层面上的问题。由重到轻简述如下。

1. 无数据 当一决策所必需的数据缺失时,勉强决策是危险的。如果时间许可,用户可先获取数据。但在许多情况下,如果只是在急需数据的情况下才去收集,很可能已影响决策,况且,并非所有的信息都是可以随时获取的。按照最顺利的情况,“一事一收集”方式会形成许多覆盖不完整的信息,相互间的标准也难统一,往往是用完即完,以后别的用户很难有效使用这些信息。在最坏的情况下,如应付突发事件(洪水淹没、旱灾等),如果没有信息贮备,而必须作出决策,就很难保证决策是明智的。就一国而言,解决这一问题的关键是对基本国土数据的需求作出预测,并制定战略规划保证适时获取和更新,最大限度地解决信息获取的无计划问题。

2. 有数据,但用户不知道数据在何处 我国信息产业发展较晚,信息资源贫乏,且往往被有关部门所垄断;即使不垄断,信息持有人也很少向潜在用户发布有关数据的信息,更不用说数据本身。在最坏情况下,用户不知已有自己想要的信息,不得已亲自收集;有时用户明知另一用户也正在收集同样的数据,但仍不敢确保自己能使用这些数据,所以仍是重复作业。过去由于国家级的机构在职能上

的交叉重复,形成同种信息“你有我有大家都有”,但互相矛盾的情况。在本次国务院机构改革中,设立“国土资源部”减少了交叉的可能性,但不能排除仍有重复的可能性,如农业部的耕地数据,水利部的水土保持统计中也包括耕地数据。如果有关各部能制定长远规划,并告之于各有关用户(其余部、省、科研院校),使任一用户都知道“过去已积累了哪些数据,现在正在收集哪些数据,将来何时会获得哪些数据”,信息共享程度将会大为提高。

3. 用户已知数据源,但无法获取数据 下列因素都会引起这一问题。

(1)费用太高 国土资源信息是国家投资后获得的,当属国家财富,但又往往为一些专业部门所垄断。受经济利益驱使,某些数据定价过高,用户无法承受。

(2)责任太重 有些数据的使用涉及到许多虽非泄密,但对数据持有人不利的因素。常见的如统计数字,有时是一种客观事实与主观意志的混合物,数据持有人知道这一背景,知道在不同场合使用不同方式解释这些数据,但他担心别的用户只从字面上理解这些数字。

(3)管理真空 谁应促进信息共享?现行的科研体制,首先鼓励用户争课题,检查用户做课题,至于成果(信息)共享,无人过问。负责成果的机构,主要精力是组织评奖,而课题评奖后,课题负责人对课题成果便不再关心了。此外,机构更迭,人事变动,会使尚未公开出版的数据变为“死数据”。

4. 已获取数据,但不知数据背景及可靠性 数据可靠性是数据质量的重要组成部分,它往往是通过背景信息提供的,如信息源、获取时间、加工方式等。如中国水土流失面积,1990年前沿用解放初期的统计数据 153 万 km^2 ,90 年代初,水利部正式发布的全国水土流失面积为 367 万 km^2 ,能否简单地讲,解放以来的 40 年,水土流失面积增加了 215 万 km^2 ?当然不是。这是因为:第一,二者覆盖的范围不同,前者没有包括新疆、西藏、台湾等六省区;第二,调查手段不同,前者使用地形图、路线调查与统计,而后者使用卫星与地面实况相结合;第三,分类标准不同,前者因当时条件制约,实际上没有数量指标,而后者有统一的土壤侵蚀分类分级标准和工作细则。本例说明,背景不清的数据,影响信息共享的质量,有时可能会得出错误的结论。但在实践中,许多数据都不提供背景信息。

5. 已获得数据及背景信息,但其格式及分类分级标准与已有数据不完全兼容 这里仅从地理信息系统对数据描述的两个要素(空间位置及属性)作简单讨论。

(1)空间位置 首先要指出,并不是因为要把数据输入计算机才强调数据标准化的重要性。任何

计算经济学初探——兼评诺贝尔经济学奖理论中的计算思想和作用

Computational Economics——On Computational Idea and Role of Theory of Nobel Prize in Economic Science

黄小原

(东北大学工商管理学院,教授 沈阳 110006)

一、什么是计算经济学

进入 90 年代,经济学领域出现了一个新的学

学科都需要共认的术语或概念,这样才能有效交流,资源科学也不例外。这里对信息系统选择何种坐标系不作讨论,只重点讨论描述资源“区域性及模糊性”特征。如果把分区、分类、分级等都看作一种广义的“术语”,信息共享的基础就是“标准化的术语”。“一事一分区”的弊端前面已作了讨论,制定全国统一的分区越早越好,尽管不一定划分太细。试设想,如果 2020 年或 2050 年后人们又重新区划出几个黄土高原,每次都在不同范围内公布一个统计数,如何谈得上连续监测?

(2)属性 通常用分类或分级的方法来描述某一属性。例如对耕地坡度的描述,是把耕地按坡度大小分成不同组。中国科学院黄土高原综合科学考察队把坡度分为 5 级($<3^{\circ}$, $3^{\circ}\sim 7^{\circ}$, $7^{\circ}\sim 15^{\circ}$, $15^{\circ}\sim 25^{\circ}$, $>25^{\circ}$);全国第二次土壤普查把坡度分为 3 级($<8^{\circ}$, $8^{\circ}\sim 25^{\circ}$, $>25^{\circ}$);水利部行业标准把坡度分为 6 级($<5^{\circ}$, $5^{\circ}\sim 8^{\circ}$, $8^{\circ}\sim 15^{\circ}$, $15^{\circ}\sim 25^{\circ}$, $25^{\circ}\sim 35^{\circ}$, $>35^{\circ}$)。当数据覆盖不完整而需综合各种数据时,这种“一事一分级”的方式不仅增加数据共享的费用,而且迫使某些数据精度退化。

促进国土资源信息共享的几点建议

标准化是实现信息共享的前提。标准化要求有统一的“名词术语”、“文件格式”、“分类编码”和“转换格式”等一系列标准,但它的基础仍是各专业领域的“名词术语”。实际上,1983~1995 年间,我国已出版《中国国家标准汇编》205 卷,涉及到信息产业的“计算机软硬件”、“数据安全”已有近百个标准。但是,各专业领域的标准化却远远滞后。如中华人

民共和国水利部于 1997 年 2 月 13 日发布强制性标准:土壤侵蚀分类分级标准(SL190—96),对术语、分类和分级标准作了规定,但其区划因数据不足仍不甚具体。作为初创期的资源科学及资源的行政主管部门,更应对国土资源信息标准化给予重视。

1. 预测国家对国土资源基本信息的需求,以此制订统一的长期规划,并让这一规划落实到具体的大型科研项目上,以时间表方式告知有关部门、省区、科研院校等。规划一旦生效,任一跨省区甚至跨县级的大型项目都要在统一的框架下使用和提供标准化的数据。

2. 促进信息交流,让更多的用户了解有关数据的信息。这既包括已有数据和正在收集的信息,还包括将来可能发布的数据。先进的计算机网络技术已经为信息共享提供了技术上的保证,但更重要的是,信息管理人员要有促进信息共享的开放意识。

3. 明确国土资源信息持有人的权利和义务。在公益性信息尚无法商品化时,找出一条让用户和信息管理者都能接受的道路,凡属收费服务,应按国家规定公开收费标准。

4. 发布信息时,应连同信息背景一同发布。

5. 国土资源信息的标准化。要总结已有成果,首先制定“推荐性行业标准”逐渐过渡到“强制性行业标准”。理想的行业标准要对国土资源各专业领域的术语、区划、分类及分级标准作出统一的规定,并且要通过各种媒介、项目的课题管理等手段使标准深入人心。对某一标准有异议和改进意义时,要专门提出建议,以备以后改进。

(责任编辑 刘先曙)