

地理数据共享效应的评价方法与应用

黄丽¹, 周卫¹, 王芙蓉²

(1. 南京师范大学虚拟地理环境教育部重点实验室, 南京 210046;

2. 南京市城市规划编制研究中心, 南京 210029)

摘要: 地理数据的信息共享具有重要性、基础性与公益性等特征。本文针对现有地理数据共享效应的评价研究较为薄弱的现状, 提出一种面向用户的地理数据共享评价策略。该方法设计了影响共享评价的指标因子, 以共享范围、共享质量、社会效益、经济效益、数据安全性, 以及共享服务作为一级因子, 并在一级因子下建立二级因子。鉴此, 分别采用层次分析法、专家咨询法构造判别矩阵, 计算因子权重, 利用缺陷扣分法与指标得分标准化的方法计算因子得分, 同时采用加权求和法计算单用户共享评价得分, 并综合各用户的评价得分获得该共享评价结果。本文以政务版电子地图数据为例, 运用本方法进行数据共享评价, 并编程实现了该功能。本方法有助于提高数据共享质量, 促进数据共享。

关键词: 数据共享; 评价; 因子; 层次分析法

DOI: 10.3724/SP.J.1047.2011.00617

1 引言

近年来, 国内外对地理数据的共享问题展开了深入探讨^[1-6]。面对数据共享的发展趋势, 评价数据共享的效果, 可提高数据共享的优劣程度, 并更好地开展数据共享工作^[7-10]。尽管各国对数据共享十分重视, 但是, 面向数据共享的评价方法、评价模型和体系是一个较新的研究课题^[1-2, 11-12]。本文在住房和城乡建设部数字化示范项目的支持下, 以政务版电子地图数据为例, 提出一种评价数据共享质量的方法。从数据共享范围、共享质量、社会效益、经济效益、安全性等方面设计评价因子, 采用层次分析法结合专家咨询法计算因子权重, 采用缺陷扣分法和因子得分标准化的方法计算因子得分, 以及采用加权求和法计算单用户共享评价得分, 并对所有用户的评价结果综合, 对政务版电子地图进行共享评价。

2 评价方法的设计与计算分析

数据共享决定了数据的使用面向不同应用部

门, 而不同部门对数据的内容、精度等有不同的需求, 因此, 为了使数据共享的评价方法合理, 本文采用面向用户的评价方法: 首先, 确定影响数据共享的评价因子, 然后, 以层次分析法结合专家咨询法计算各因子的权重, 进而让用户对该共享数据进行打分, 最后, 对所有用户的评价结果进行综合, 得到该政务版电子地图数据共享的最终得分。

2.1 评价因子与指标设置

根据政务版电子地图数据共享的实际情况, 本文选择共享范围、共享质量、社会效益、经济效益、数据安全性, 以及共享服务作为一级评价指标, 再在每个一级指标的基础上设置二级指标, 进行数据共享的评价。具体指标的设置如表1所示。

2.2 因子权重计算与评价层次分析

本文采用层次分析法、专家咨询法计算了评价因子的权重。如图1所示, 分别为A—B_i(数据共享评价层一级因子重要性比较), B₁—C_{1j}(共享范围层二级因子重要性比较), B₂—C_{2j}(共享质量层二级因子重要性比较), B₃—C_{3j}(社会效益层二级因子重要

收稿日期: 2010-09-02; 修回日期: 2011-09-19.

基金项目: 国家“863”计划项目(2006AA12Z222)。

作者简介: 黄丽(1986-), 女, 江苏扬州人, 硕士, 主要从事基础地理数据共享、GIS版权保护等方面研究。

E-mail: huangli_gis@163.com

性比较), $B_4 - C_{4j}$ (经济效益层二级因子重要性比较), $B_5 - C_{5j}$ (数据安全性层二级因子重要性比较), $B_6 - C_{6j}$ (共享服务层二级因子重要性比较) 构造判断矩阵, 从而计算出一级因子的权重 B_i , 二级因子

相对于其一级因子的权重 C_{ij} , 进而计算得到各二级因子在数据共享评价中的总权重 $W_{ij} = C_{ij} * B_i$, 其中, $i, j = 1, 2, \dots$ 。

表 1 政务版电子地图数据共享的评价因子

Tab. 1 Assessment factors of government digital map data sharing

一级因子	二级因子	因子概念
共享范围	数据共享率	使用共享数据的用户的数量占政府部门数量的百分比
	用户影响力	使用共享数据的用户对社会进步、经济发展的影响力
共享质量	数据数学精度	共享数据的数学精度评价主要包括数据采用的坐标系、投影系、高程系是否正确, 平面位置精度、高程精度、接边精度是否符合要求等
	数据属性精度	共享数据的属性精度评价, 主要包括要素分类与代码是否正确, 要素属性值是否正确, 属性项类型是否完备, 数据分层是否正确完备, 注记是否正确等
	数据逻辑一致性	共享数据的逻辑一致性检查, 主要包括应建立拓扑关系的要素层和要素是否有遗漏, 拓扑关系是否正确, 多边形是否闭合, 结点是否匹配等
社会效益	数据完备性	共享数据的要素完备性检查, 主要包括数据的内容是否完整, 有无遗漏和重复; 要素层、要素项是否有遗漏, 是否缺少或重复; 注记是否完整等
	效率的提高	用户的工作效率在使用共享数据后是否得到提高
经济效益	社会影响力	用户使用共享数据的所有项目对社会产生的总影响力
	效益增加	用户在使用共享数据后增加的直接和间接经济效益
数据安全性	法律法规要求	共享数据是否遵守国家安全、知识产权、隐私权等相关法律法规
	保密手续	数据提供部门与用户之间是否签订保密协议, 保密手续办理是否完备
	保密技术手段	数据提供部门对共享数据的保密技术手段是否合理, 对用户的权限设置是否满足用户需求
共享服务	数据适用性	共享数据是否满足用户的需求, 若不满足需求, 用户还需投入的对数据加工与处理, 直至满足需求的工作量, 投入的工作量越小说明数据适用性越好。反之, 数据适用性越差
	技术支持能力	针对用户提出的问题是能及时响应并提供相应的解决办法
	服务质量满意度	用户对数据提供部门服务质量的满意度

其中, 层次分析法的具体实现步骤如下^[13]:

(1) 用层次分析法, 依据指标间的相对重要性构造判断矩阵 $A = a_{ij}$ 。

其中, a_{ij} 为第 i 个指标相对于第 j 个指标的重要程度取值, 一般具有如下特征: $a_{ij} > 0, a_{ij} = 1/a_{ji}, a_{ii} = 1, (i, j = 1, 2, \dots, n)$ 。

(2) 依据判断矩阵, 采用几何平均法(见公式 1) 计算权重向量 W , 以及相应的最大特征值 λ_{max} (见公式 2)。

几何平均法公式表示为:

$$W = \frac{\left(\prod_{j=1}^n a_{ij}\right)^{1/n}}{\sum_{i=1}^n \left(\prod_{j=1}^n a_{ij}\right)^{1/n}} \quad (1)$$

最大特征值 λ_{max} 的计算公式为:

$$\lambda_{max} = \frac{1}{n} \sum_{i=1}^n \frac{Aw_i}{w_i} \quad (2)$$

(3) 一致性检验。检验依据为, 计算一致性指

标 $C. I.$ (见公式 3); 查表得随机平均一致性指标 $R. I.$ (表 2), 并计算一致性比率 $C. R.$ (见公式 4)。若 $C. R. \leq 0.1$, 则认为评判的过程具有满意的一致性; 否则要对矩阵中的元素作以调整, 使之具有满意的一致性。

一致性指标 $C. I.$ 的计算公式为:

$$C. I. = \frac{\lambda_{max} - n}{n - 1} \quad (3)$$

一致性比例 $C. R.$ 的计算公式为:

$$C. R. = C. I. / R. I. \quad (4)$$

表 2 随机平均一致性指标 $R. I.$

Tab. 2 Mean random consistency index $R. I.$

阶数	1	2	3	4	5	6	7	8	...
$R. I.$	0	0	0.52	0.89	1.12	1.26	1.36	1.41	...

2.3 因子得分计算

根据各评价因子的不同, 分别采用缺陷扣分

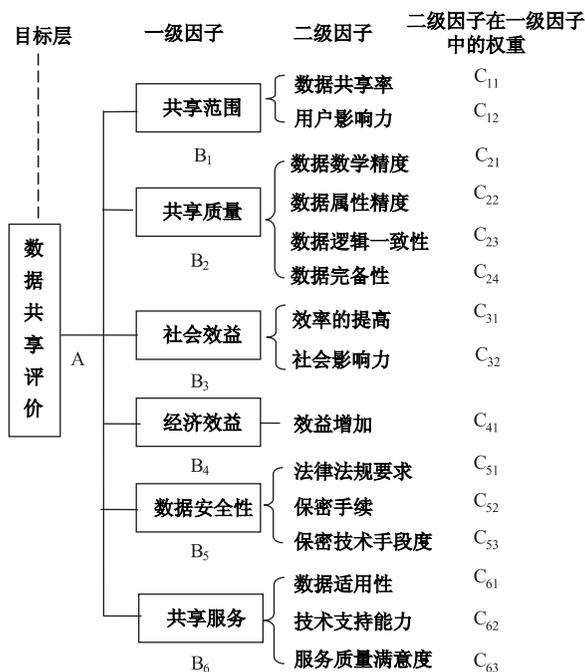


图1 数据共享评价层次分析图

Fig.1 Hierarchy analysis graph of data sharing assessment

法,以及评价指标得分标准化的方法,对各指标进行得分评定。

(1) 针对共享质量、数据安全性、共享服务 3 类因子的特性,本文采用缺陷扣分法进行该类别指标的得分评定。评分准则为:同一问题不重复扣分,每项最低得分为 0 分。严重缺陷扣 42 分,重缺陷扣 12 分,次重缺陷扣 4 分,轻缺陷扣 1 分^[14]。计算公式如下:

$$X = 100 - 42 * n1 - 12 * n2 - 4 * n3 - n4 \quad (5)$$

其中, $n1$ 、 $n2$ 、 $n3$ 、 $n4$ 分别为严重缺陷、重缺陷、次重缺陷和轻缺陷的个数。

(2) 针对共享范围、社会效益、经济效益这 3 大类因子的特性,本文将因子得分标准化的方法进行因子得分判定^[15]。由于这 3 类因子都是正相关的,因此,可以采用公式 6 进行标准化:

$$X = y/Y * 100 \quad (6)$$

其中, X 为某因子的得分, y 为该因子的实际值, Y 为该因子的标准值。当该因子的实际值大于标准值时,因子得分取 100 分,其打分方法:

(1) 共享范围因子打分方法

① 数据共享率:其得分可以通过用户数量计算得到。本文设定 15 为数据共享率的标准值。

② 用户影响力:不同的数据共享部门对于社会的影响力大小不同。综合考虑现有关于政府部门影响力的分类研究成果和各政府部门对数据共享的影响力的实际情况,对各政府部门的影响力进行分组并给出每组的得分范围。本文设定 12 为用户影响力的标准值。

具体分组情况如表 3 所示。

表 3 用户影响力分类得分表

Tab. 3 Classifying and scoring by user's influence

组别	政府部门	分值
强影响力组	公安部门、国土部门、房产部门	12-15
中等偏强影响力组	测绘部门、城乡规划部门、物流部门	9-12
中等强度影响力组	民政部门、环保部门、旅游部门	4-8
弱影响力组	教育部门、工商管理部、发展改革部门等其他部门	1-3

(2) 社会效益因子打分方法

① 效率的提高:假设用户在使用该共享数据前完成某工作量的时间为 t ,使用该共享数据后减少的工作时间为 t' ,则效率提高的比率为 t'/t 。本文设定 1/2 为效率的提高的标准值。

② 社会影响力:用户使用部门在使用该共享数据后,对社会影响力的提升产生的作用。具体是指共享数据的使用对居民生活、生态环境、社会环境、公共服务四个方面的提升效果,如果共享数据的使用直接服务于以上四个方面中的一个或多个,均视为社会影响力的提高。本文设定社会影响力提高的标准值为 2,即 2 个方面。

(3) 经济效益因子打分方法

假设用户在使用该共享数据前的效益为 p ,使用该共享数据后增加的效益为 p' ,则效益增加的比率为 p'/p 。本文设定 1/2 为效益增加的标准值。

3 数据共享评价的应用实例

本文根据上述数据共享的评价方法,以住房和城乡建设部城市数字化工程示范项目中政务版电子地图为例,对其进行了数据共享的评价(流程如图 2 所示)。

根据图 2 的评价流程,本文对各评价因子进行权重计算,采用层次分析法结合专家咨询法得到各

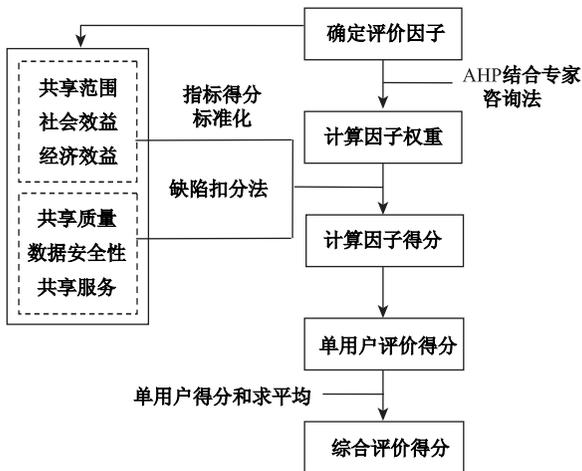


图2 政务版电子地图数据共享评价流程图

Fig. 2 Flowchart of government digital map data sharing assessment

评价因子的权重如图3所示。

单用户得分评价采用缺陷扣分法,以及评价指标得分标准化的方法进行得分评定。再结合各因子权重,采用加权求和法求得各用户对该政务版电子地图数据共享的评价得分。根据这一方法可以得到12个用户的评价得分分别为:84.24、86、79.5、88.2、74.92、70.6、85、90.5、77.4、84.3、88,该市政务版电子地图数据共享评价总体得分= $(84.24+86+79.5+88.2+74.92+70.6+85+90.5+77.4+84.3+88)/12=83.31$ 。

在上述评价方法设计的基础上,本文基于.net平台采用c#语言开发实现了共享评价的主要功能,包括因子权重计算(图4(a))与单用户得分计算(图4(b))。

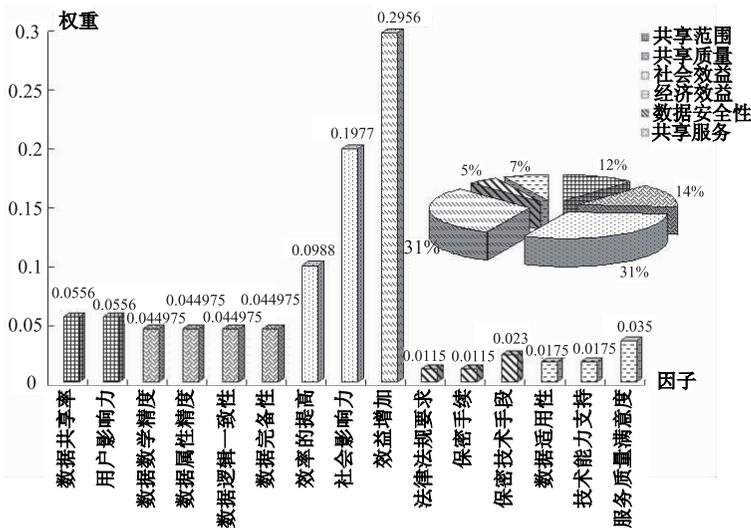


图3 各因子权重

Fig. 3 Weight of each factor



图4 系统界面

Fig. 4 System interface

4 结语

针对现有数据共享评价研究较为薄弱的现状,以政务版电子地图数据为例,设计了一种面向数据使用者的数据共享成效的评价方法。该方法能够较客观地反映数据共享情况,对于提高数据共享的质量有直接的促进功能,研究可为更好地进行数据共享工作提供依据。该方法中涉及的指标因子体系、缺陷扣分标准与因子得分标准值,主要是针对政务版电子地图数据的共享评价设计,在对其他类别的空间数据进行共享评价应用时,可以根据自身数据的特点等,调整指标体系、缺陷扣分标准与因子得分标准值,以满足自身数据要求的共享成效评价。

参考文献:

- [1] 何建邦,柯正谊,吴平生,等. 国家地理信息共享政策框架意见[J]. 资源科学, 2001, 23(1): 17-22.
- [2] 孙枢,马宗晋,周秀骥,等. 我国地球科学数据共享问题[M]. 北京: 科学出版社, 1998.
- [3] Dueker K J, Butler J A. A Geographic Information System Framework for Transportation Data Sharing[J]. Transportation Research Part C: Emerging Technologies, 2000, 8(1-6): 13-36.
- [4] Keller S F. Modeling and Sharing Geographic Data with INTERLIS[J]. Computers & Geosciences, 1999, 25(1): 49-59.
- [5] Frank S. Cataloging Digital Geographic Data in the Information Infrastructure: A Literature and Technology Review[J]. Information Processing & Management, 1994, 30(5): 587-606.
- [6] 王连备,贲进,吴英. 面向服务的空间数据共享技术研究[J]. 测绘通报, 2010, (7): 68-70.
- [7] 何建邦,闻国年,吴平生,等. 地理信息共享法研究[M]. 北京: 科学出版社, 2000.
- [8] 陈腾. 国土资源科学数据共享平台的数据发布技术研究[J]. 河南理工大学学报(自然科学版), 2010, 29(1): 66-69.
- [9] Fonseca F T, Egenhofer M J, Davis Jr. C A, Borges K A V. Ontologies and Knowledge Sharing in Urban GIS[J]. Computers, Environment and Urban Systems, 2000, 24(3): 251-272.
- [10] Mennis J, Guo D S. Spatial Data Mining and Geographic Knowledge Discovery—An Introduction[J]. Computers, Environment and Urban Systems, 2009, 33(6): 403-408.
- [11] Pundt H, Bishr Y. Domain Ontologies for Data Sharing—An Example from Environmental Monitoring Using Field GIS[J]. Computers & Geosciences, 2002, 28(1): 95-102.
- [12] 樊文有,谢琼,王永东. GIS空间数据共享交换平台的设计[J]. 测绘通报, 2010(6): 47-50.
- [13] 方新华. 小城市投资环境评价体系研究[D]. 浙江大学硕士论文, 2004.
- [14] 国家质量技术监督局. GB/T18316-2001, 数字测绘产品检查验收规定和质量评定[S]. 北京: 中国标准出版社, 2001.
- [15] 陈强. 南京市栖霞区土地集约利用研究[C]. 2008年中国土地学会学术年会论文集, 2008, 1355-1362.

A Method and Application of Assessing the Effort of Geographic Data Sharing

HUANG Li¹, ZHOU Wei¹, WANG Furong²

(1. Key Laboratory of Virtual Geographic Environment, Minister of Education, Nanjing Normal University, Nanjing 210046, China; 2. Nanjing Urban Planning and Research Center, Nanjing 210029, China)

Abstract: With the high costs of geographic data and its conversion the data transfer barrier becomes a major obstacle in reusing geographic data for geographical information systems (GIS). Data sharing is one of the most effective methods on breaking this barrier which can promote developing and reusing geographical data. It has broken the monopoly of information resources and reduced the "digital divide". So it has great significance in promoting social and economic development. For the lack of researches in data sharing assessment, we proposed a user-oriented method of assessing the data sharing. For the same sharing data, the results of data sharing assessment may be different with different users. So when assessing a sharing

data, each user should evaluate firstly and then all results should be considered in the final evaluation result. We firstly design the assessing factors which affect data sharing. In this paper, we make range of sharing, quality of sharing, social benefit, economic benefit, data security and services of sharing as first-grade factors, and on the basis of each first-grade factors, we design corresponding second-grade factors. Secondly, we construct judgment matrixes and calculate weighs of factors by AHP (Analytic Hierarchy Process) and Expert Consultation Method. Thirdly, according to different factors, we use deflection subtraction method or method of score standardization in calculating scores of factors. Fourthly, we use the method of weighted sum to calculate the result of data sharing assessment of each user. After getting the results of all users, the average value is calculated as the final score of the data sharing assessment. In this paper, we use the method proposed to assess the data sharing of the government digital map as an example and the functions are implemented by coding. The method realizes the assessment of data sharing, which is benefit in improving the quality of data sharing. It can promote the performance of data sharing and reduce the wasting of resources in data sharing.

Key words: data sharing; assessment; factor; AHP