

基于知识库的图书馆绩效评价与 辅助决策系统的设计与实现^{*}

□郭向勇 陈锐浩 敖龙 李勇 汪勇 龚自振

摘要 提出基于图书馆行业知识库的绩效评价和辅助决策系统的总体设计与实现方法,并重点介绍支撑评价系统的两层分析模型。第一层分析,以用户录入的图书馆事实数据为依据,通过分析得分规则库,实现智能打分;第二层分析,根据各指标得分情况,通过匹配算法连接知识数据库,推导出评价信息和指导性建议,并通过案例验证设计方法的实现过程。

关键词 图书馆 绩效评价 知识库 智能分析 辅助决策

1 引言

随着高等职业教育的蓬勃发展,高职院校对图书馆事业的建设和发展寄予了更大期待。通过建立高职院校文献信息资源保障体系的绩效评价体系来促进图书馆事业的建设和发展,已经引起广泛关注。先后有全国或部分省市高职院校图书馆建设指南与评价指标体系出台^[1-2]。今年广东省教育厅正式发布广东省高校图书情报工作指导委员会高职分委员会编制的《广东省高职院校图书馆评价指标体系》(试行),并委托深圳职业技术学院开发广东省高职院校图书馆建设评价与辅助决策系统平台。该平台以各馆事实数据为基础,实现由事实数据向评价指标的智能转换,通过评价指标统计进行二次分析,给各图书馆提供建设现状的客观统计分析结果与绩效评价,发挥辅助决策的功能。

本文主要论述系统设计目标、模型、模块设计和关键技术等。

2 需求及技术思路

2.1 国内外研究现状

关于图书馆建设与发展的绩效评价与辅助决策系统,具有代表性的中外相关文献有:姜合等的基于数据仓库的高校图书馆决策支持系统设计,偏重为

图书流通和图书采购等事务决策提供依据^[3];王志强的关于高校图书馆文献采访模糊决策支持系统,侧重于订购需求库设计^[4];胡华南等的关于高校图书馆书刊采访决策支持系统,侧重建立采购经费分配和书刊选购决策模型^[5];夏勃等提出关于 Web 荐购图书采访决策支持系统设计方法^[6];李默提出高校图书馆图书荐购系统决策支持模型^[7];陈晓等提出基于业务数据的高校图书馆决策支持系统设计方案^[8]等。Robin 提出一种基于管理者需求分析的图书馆商业决策支持系统^[9];Pérez, I. J. 通过构建决策支持系统模型来解决动态群体决策问题的移动决策支持系统^[10]等。

2.2 需求分析与设计思路

鉴于我国高等职业教育发展还处于初创阶段,高职院校图书馆的整体建设与发展水平参差不齐,经费不足、专业队伍缺乏和服务方式单一等是高职院校图书馆存在的共性问题。通过绩效评价方式整体性地推动与促进图书馆事业的健康与有序地发展,开发基于知识库的绩效评价与辅助决策系统是最佳选择。

系统知识库由图书馆事实数据、行业专家提供的评价信息和建设指导意见,以及形成这些内容的各种规则和推导方法等内容组成。其中评价信息和

^{*} 本文系教育部人文社会科学规划基金项目“中国高职院校图书馆协同管理评价体系的构建与实践研究”(项目编号:12YJA870006)的研究成果之一和广东省教育厅 2010 年度教育教学改革项目“广东省高职院校图书馆建设评价与辅助决策系统”(项目编号:GZZZD2011005)的研究成果。

建设指导内容是根据广东省高职院校图书馆评价指标体系形成的,该指标体系共有一级指标 9 项、二级指标 24 项、三级指标 75 项和评分细则 236 项。对应各级各项指标和细则,按照不同的评价结果,由图书馆专家经过多轮次讨论形成和阶段性固化,最终确定相应的绩效评价信息、辅助决策意见和建议。

系统采用基于二维表的知识表达方法,把知识内容和规则以二维表方式存储在数据库表中,二维表建立知识属性与值间的对应关系,而且便于动态维护和管理。

知识形成过程如图 1 所示。系统通过集中数据,实现馆际纵向和横向对比与分析。图书馆用户录入系统的事实数据经过分析和比较形成评价指标体系的绩效考核信息,通过进一步匹配和分析,形成最终的知识,即绩效评价和辅助决策的意见和建议。

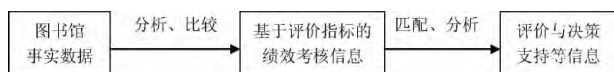


图 1 知识形成过程示意图

3 具体的实现方案

3.1 系统模型

系统模型示意图如图 2 所示。系统底层数据包括事实数据和调查数据,事实数据和调查数据由各图书馆填写,根据本馆实际情况,填写各种评价指标对应的调查表。

系统直接对这些原始数据进行两层分析。第一层分析,根据评价指标模型和细则得分规则,把事实数据和调查数据自动转换成各评价指标的得分结果。第二层分析,根据各指标具体得分,采用区间匹配算法,从知识库获得对应的绩效评价和辅助决策信息,最后以报表的形式呈现给图书馆用户。

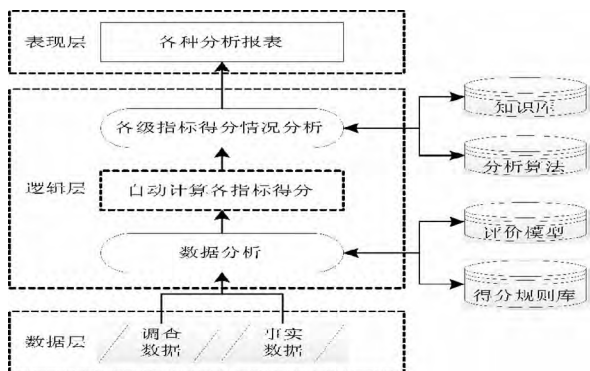


图 2 系统平台模型示意图

3.2 模块设计

系统依据《广东省高职高专院校图书馆建设指南与评价指标体系》和《广东省高职高专院校图书馆评价指标和辅助决策细则》2 个文档构筑系统平台知识库核心内容,并根据用户角色和系统数据流确定具体功能模块。

具体角色有管理员、图书馆用户、图书馆馆长、省厅用户和游客。

系统分为前台部分和后台部分,前台部分主要面向图书馆用户和省厅用户。

图书馆用户主要功能模块有:数据填报、统计查询、绩效评价、辅助决策和同行比较等。

省厅用户主要功能模块有:信息查询、统计分析和图书馆比较等。

后台部分是面向管理员的功能模块。包括系统管理、知识管理和统计分析等主要模块。管理员通过系统管理模块可进行系统各种参数的配置,在知识管理模块可动态维护和更新评价指标体系,以及定义各种得分规则、匹配规则和维护图书馆绩效评价和辅助决策信息。

3.3 系统关键技术

(1) 多种不同自测题目的自动打分功能。

事实数据和图书馆自我评价数据是系统对图书馆评价的依据,这些数据需要图书馆用户录入。为了能够给出细化评价精度,系统设置 9 项一级评价指标,24 项二级指标、75 项三级指标,每个指标栏目的评价内容包含以评价指标细则为数据框架的多种填报类型。填报类型有逻辑型、单选型、多选型、数字型和计算型等。系统根据评分规则自动实现用户数据到具体分值的转换。

逻辑型只有 Yes 和 No 两个选项,Yes 选项才能得分;单选型,每个选项都能得分,而且分值不一样,每个选项的具体分值存储在细则数据表里,每个选项对应一个分值字段,具体分值在系统后台设置。多选型,用户可以选择多个选项,每个选项的分值可以相同,也可不同,分值为多个选项的得分累加。单选题目的得分计算直接采用字符串匹配,多选采用 C# 语言的 Indexof() 函数进行判断。

数字型得分规则根据后台定义的关联,具体数据从用户填报的事实数据中自动获取,系统采用二维表存放具体规则,通过遍历规则表映射提交数据和具体规则的关系,把用户提交的数据转换为具体

的得分值。

数字型得分规则表为:把整数区间细分为若干个区间,每个区间分数配置不一样,判断用户提交的数值处于哪个区间,具体得分就是所在区间对应的分值。表1为规则表数据结构,表2为生均数字咨询参考业务量在数据库中的得分规则。

表1 规则表数据结构

序号	列名	数据类型	长度	小数位	标识	主键	允许空	说明
1	ID	int	4	0	是	是	否	
2	ZBID	char	10	0			是	指标ID
3	Node	nvarchar	50	0			是	指标细则
4	LowerNum	float	8	0			是	变量左区间范围数值
5	Left Operation	nvarchar	50	0			是	左边计算规则
6	Right Operation	nvarchar	50	0			是	右边计算规则
7	Height Num	float	8	0			是	变量右区间范围数值
8	Score	float	4	0			是	处于该区间的得分

表2 生均数字咨询参考业务量在数据库中的得分规则

ZBID	Node	Lower Num	Left Operation	Right Operation	Height Num	Score
1	C343	0.01	=<			4.0
1	C343	0.001	=<	<	0.01	3.0
1	C343	0.0001	=<	<	0.001	1.0
1	C343	0.0	=<	<	0.0001	0.0

规则内容是根据专家讨论后形成的,支持动态更新。在遍历规则表时,若已找到所处区间,就可提前跳出循环,返回用户呈现细则得分,并存放在变量Score中。

数字型数据转换为具体得分的程序流程图如图3所示。

计算型得分则根据用户提交的若干数据,选择一种具体的计算方法,先计算出具体的数据,然后按照数字型得分予以处理。

(2) 建立知识库和相应规则的映射实现智能评价和辅助决策。

为有效指导图书馆的建设,系统依据《广东省高职院校图书馆评价指标体系和细则》,结合图书馆专家意见、图书馆基础信息和填报数据,建立完整的绩

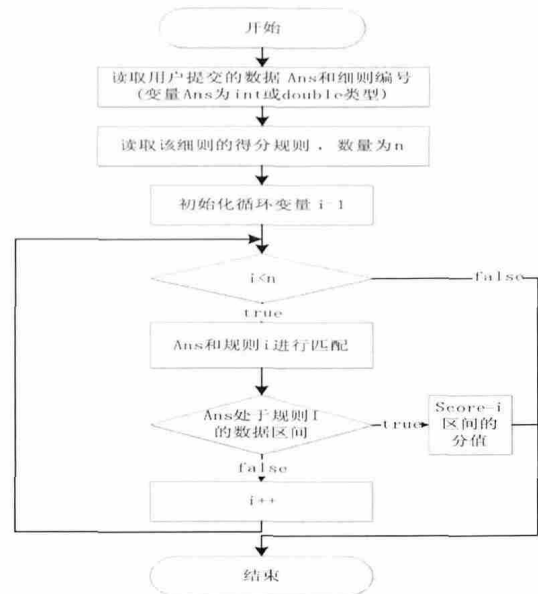


图3 数字型细则计算得分流程图

效评价知识库和辅助决策知识库。知识库支持动态管理,通过对具体指标细则的得分进行分析,实现从指标得分到知识库的智能映射,自动形成评价信息。

智能评价算法的实现分预处理和匹配两个阶段:预处理阶段,根据传递评价指标栏目参数,构造匹配串S,并且从知识库中找出相关的评价信息集;匹配阶段,根据得分结果,遍历规则库,寻找所属的区间。如果找到,则读取该区间对应的评价信息,匹配成功。

评价信息知识库具体数据格式,如表3所示。

为了对算法进行描述,定义如下符号和函数:

评价信息知识库数据集为: $A = \{R_i | i = 1, 2, \dots, n\}$, n 为数据集总数。

匹配串: $S = \{Node_j | j = 1, 2, \dots, m\}$, m 为 Node 栏目评价细则数量。

表3 评价信息知识库数据结构

序号	列名	数据类型	长度	小数位	标识	主键	允许空	说明
1	ID	int	4	0	是	是	否	
2	ZBID	int	4	0			是	当前使用评价指标体系,外键
3	Node	nvarchar	50	0			是	指标细则编号,外键
4	LowScore	float	8	0			是	得分范围下限
5	Height Score	float	8	0			是	得分范围上限
6	Content	ntext	4	0			是	专家评价信息资料

具体算法描述如下:

第一阶段:预处理阶段,包括下面两个步骤。

(1)读取用户传递的指标参数,即指标栏目编号,记为 $Node$ 。接着读取所有属于该栏目的评价细则,取出细则编号构成匹配串 S 。

(2)以 S 为检索条件,从知识库中检索出所有跟 S 有关联的记录,记为集合 $B, B \in A$ 。

第二阶段:匹配阶段,由于经过预处理,得到的记录集 B 已经是最精简的,所以采用遍历方式匹配,具体步骤如下:

(1)初始化循环变量 $j=1$;

(2)取出 S 中第 j 个细则编号 $Node_i$;

(3)执行 $Find(Node_i, B)$, 遍历集合 B , 如果匹配成功,返回该细则评价结果;

(4)循环变量 $j++$, 如果 $j>m$, 结束循环,否则跳转至 2;

匹配函数 $Find(Node, B)$ 参数说明: $Node$, 指标细则编号, 系统支持多套指标, 多个年份, 多个用户, 由于指标、年份、用户这三个变量是全局变量, 所以函数参数中不包含它们; B , 知识库中跟指标细则 $Node$ 有关联的记录集, 记录集长度为 k 。

$Find(Node, B)$ 算法描述如下:

(1)根据 $Node$ 参数, 加上指标、年份、用户三个全局变量, 作为检索条件从得分情况事实表中获得用户该细则的得分情况, 记为 $Score$;

(2)初始化循环变量 $i=1$;

(3)读取集合 B 中的 R_i ;

(4)如果 $Score$ 处于 R_i 定义的得分区间 $[LowScore, HeightScore)$, 读取 R_i 的 Content 字段内容并返回。读取状态设为 true。

(5)循环变量 $i++$, 如果 $i>k$, 结束循环, 否则跳到 3;

算法分析第一阶段, 主要是和数据库进行交互, 执行两次数据库查询, 第一次查询, 生成模式匹配串; 第二次查询, 得到跟模式串相关的记录集, 存放在泛型集合 $List<T>$ 中。第二阶段, 由于执行两次循环, 所以时间复杂度为 $O(mk)$, m 为匹配串长度, k 为匹配串关联记录集长度。虽然采用遍历算法, 但是由于预处理阶段对知识库数据集做了筛选, k 的数值并不大, 所以运行效率还是得到保证。

4 应用效果评价

以 C34 指标填报为例予以验证。

图书馆用户登录后, 需要先填报基础数据和指标自评数据, 如图 4 所示。该图左边是整个评价指

标体系, 以目录树显示。点击目录树第三级指标“C34 数字咨询参考”, 右边显示该指标的评价细则, 其中 C341 和 C342 是单选型, C343 是计算型。计算两变量是 2011 年数字参考咨询业务和在校学生人数, 这两个变量数值是根据后台设置自动从基础数据中获取的, 两变量根据计算规则形成生均数字参考咨询业务量。

图 4 前台图书馆用户提交数据页面

系统第一层分析, 根据得分规则计算各细则的得分。图 4 中指标细则 C341 是单选型, 用户选项为 A, 系统自动给出选项 A 的分值, 具体为 4 分; 指标细则 C342 也是单选型, 用户选项为 A, 系统自动给出选项 A 的分值, 为 2 分; 指标 C343 生均数字参考咨询业务量是计算型, 系统自动获取该图书馆用户填写的事实数据“年数字参考咨询业务量”为 2300 次, 在校学生人数为 22300 人, 根据定义的计算规则, 自动生成“生均数字参考咨询业务量”为 0.1031 次/人, 该数据根据图 3 介绍的流程智能转换为得分。通过遍历算法, 根据得分规则找到匹配记录, 如表 2 所示。由于 0.1031 处于 $[0.01, +\infty)$, 第一条记录满足条件, 所以自动判断用户该细则得分为该规则的分值, 即 4 分。

第二层分析, 实现智能评价和给出指导性意见和建议, 以指标 C34 为例, 第一层分析中已经自动计算出各细则具体得分, 第二层分析预处理阶段, 用户传递指标参数 $Node=C34$, 程序会执行 SQL 语句, 查找所有父栏目结点为 C34 的评价细则, 检索出来之后构成匹配串 $\{C341, C342, C343\}$ 。同时执行 SQL 语句, 从评价信息知识库中检索数据, 用泛型集合 $List<T>$ 存储, 具体 SQL 为:

```
Select * from 评价信息表 where Node in ('C341'
'C342' 'C343') and ZbID=当前使用指标体系。
```

匹配阶段, 先读取 C341, 执行 $Find(C341, List<T>)$ 得到 C341 的评价信息, 接着是 C342, 最后是 C343, 遍历完成之后, 返回整个 C34 栏目的评价

信息。如图 5 所示。

2011 年度栏目“C34 数字参考咨询”分析评价内容

2009年 2010年 2011年					
参评图书馆数	满分	我的得分	平均分	最高分	最低分
56 所	10	10	6.96	10	0

分析评价:

- C341. 数字参考咨询服务提供主动服务和个性化服务, 形式多样
- C342. 数字参考咨询服务效果良好, 满足读者在信息服务方面的需求
- C343. 主均年数字参考咨询业务量较多, 较好地开展了面向学生读者的参考咨询服务工作

图 5 指标 C34 评价结果

2012 年 5 月, 该系统平台上线运行使用, 收效甚好。2012 年 7 月经深圳市科技工贸和信息化委员会组织包括图书馆界专家在内的专家组进行科技成果鉴定, 专家们一致认为: “该系统平台有创新、实用性强、深入研究图书馆建设指标体系, 并转换为实时、动态的基于智能分析的计算机软件系统, 在全国图书馆界首次采用, 居国内领先水平, 具有推广应用价值”。

5 结论

基于知识库的图书馆绩效评价与辅助决策系统以《广东省高职高专院校图书馆评价指标体系和细则》为依据, 在实践的过程中不断更新与发展知识库的知识信息, 从而确保系统平台的科学性、实用性、时效性和可操作性。

评价指标与辅助决策细则体系有效地构筑了该系统平台知识库的核心内容, 这是构建系统平台的关键所在。同时系统平台首次提出了对图书馆基础数据的两层分析与应用研究, 并通过系统平台运行予以实践验证。

该系统不仅是一个评价系统, 更是一个学习系统, 不但能发现问题, 还能解决问题, 这是有别于任何评价系统的特色所在。正因为如此, 特别适合于高职院校图书馆建设与发展的需求, 能够很好地发挥引导作用。

参考文献

- 教育部高等学校图书馆情报工作指导委员会. [2013-02-26]. <http://www.scal.edu.cn/courseInfoSearch.html?miscdictId=4>
- 广东省教育厅. [2013-02-26]. <http://www.gdhed.edu.cn/main/search.php>
- 姜合, 栾秀梅, 张学云. 基于数据仓库的高校图书馆决策支持系统设计. 图书馆理论与实践, 2005, (8(4)): 81-83
- 王志强. 高校图书馆文献采访模糊决策支持系统及应用. 湘潭师范学院学报(自然科学版), 2009, 31(1): 163-167
- 胡华南等. 高校图书馆书刊采访决策支持系统. 数学的实践与认识, 2009, 39(7): 8-13
- 夏勃, 白似雪. 基于 Web 荐购的图书采访决策支持系统设计. 图书馆学研究, 2012, (1): 21-24
- 李默. 高校图书馆图书荐购系统的决策支持模型研究. 情报探索, 2010, 156(10): 103-105
- 陈骁等. 基于业务数据的高校图书馆决策支持系统设计. 中国现代教育装备, 2012, 23 (159): 82-84+93
- Robin A. Paynter. Commercial Library Decision Support Systems: An Analysis Based on Collection Managers' Needs. Collection Management, 2008, 34(1): 31-47
- Pérez, I. J., Cabrerizo, F. J., Herrera-Viedma, E. A Mobile Decision Support System for Dynamic Group Decision-Making Problems. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 2010: 1244-1256

作者单位: 郭向勇, 敖龙, 李勇, 汪勇, 龚自振 深圳职业技术学院图书馆, 深圳, 518055
陈锐浩 深圳职业技术学院教育技术与信息中心, 深圳, 518055

收稿日期: 2013 年 5 月 22 日

Design and Implementation of Library Performance Evaluation and Decision-Making Support System Based on Knowledge Base

Guo Xiangyong Chen Ruihao Ao Long Li Yong Wang Yong Gong Zizhen

Abstract: For the purpose of improving library management level and advancing the operability and practicability of the performance evaluation of vocational colleges, the paper offers the scheme which is based on Library Knowledge Base Performance Evaluation and Decision-Making Support System. It introduces the implementation methods, with a strong focus on its two-stage analysis model in performance evaluation system. The first stage of analysis is using raw data entered by users, analyzing and scaling to achieve an intellectual score. The second stage is to use each intellectual score in Knowledge database via matching algorithm, to gather evaluation information and constructive suggestions. Implementation methods has then been introduced furthermore.

Keywords: Library Performance Evaluation; Knowledge Database; Intelligent Analysis; Decision Support