

文章编号: 1001-1986(2008)06-0017-03

数据字典在煤炭数据发布平台中的应用

杨圣伟, 汪新庆

(中国地质大学数学地质与遥感地质研究所, 湖北 武汉 430074)

摘要: 介绍了数据字典技术在基于网络 B/S 结构的“全国煤炭资源网络信息系统”——煤炭数据发布平台中的应用情况。系统通过数据字典技术进行用户访问控制、数据语义规格化和显示页面动态定制等工作。数据字典技术在该系统中的应用, 不仅使系统管理人员能够统一、高效、灵活地进行访问控制, 而且能根据需求定制显示内容, 保证了系统的安全性和功能的完善性。

关键词: 数据字典; 煤炭数据库; 数据发布

中图分类号: P628.4 **文献标识码:** A

Application of data dictionary in coal database system under B/S

YANG Sheng-wei, WANG Xin-qing

(Institute of Mathematical Geology and Remote Sensing Geology,
China University of Geosciences, Wuhan 430074, China)

Abstract: The paper introduced “The Coal Resource Network Information System” which based on WEB under B/S mode and the using of data dictionary in the system. The system used the data dictionary on user access control, data meaning translating, and page dynamic customization. The application of data dictionary in the system can make the system manager effectively conduct user access control, display content based on requirement, ensure system’s safety and functions.

Key words: data dictionary; WEB database; data publishing

我国地学数据的公众发布与共享一直落后于其他行业, 煤炭数据至今没有一个统一的面向大众的发布平台。为了配合新一轮的全国矿产资源潜力评价工作, 必须建立全国范围的、面向大众的煤炭数据发布系统。“全国煤炭资源网络信息系统”是基于 1992 年全国第三次煤田预测数据, 在原有“全国煤炭资源信息系统”的基础上, 根据原系统的业务模式与功能需求, 建立了基于 MS SQL Server 的大型网络数据库, 并采用 Microsoft .NET 平台, 研发出基于网络 B/S 结构的网络信息系统。该系统是全国煤炭行业首个面向大众的大型数据库系统, 它的应用使煤炭资源数据实现了跨系统、跨平台和跨地域共享。

数据字典(Data Dictionary, DD)是描述数据库中各数据属性与组成的数据集合, 有人把它看作是关于数据库中数据的数据库。数据库系统是一个复杂的系统, 其中所包含的信息除了存入的用户数据外, 还有很多非用户数据信息。例如, 模式和子模式的内容、文件间的联系、数据项的长度、类型、用户

标识符、口令、索引等等。这些信息一般都集中保存在一个专门的地方, 这就是数据字典^[1]。

数据字典技术在数据处理方面具有明显的优势。经过多年的发展, 数据字典技术已取得了很大进步, 并且在很多方面进行了卓有成效的应用。但是, 因为地学数据特有的多源、多量、多类、多元和多维等特点, 使得数据字典在地学数据发布, 特别是在煤炭数据发布的应用上还存在不足。本文将数据字典技术引入“全国煤炭资源网络信息系统”, 探讨其在煤炭数据库中的应用情况。

1 基于数据字典技术的用户权限控制

1.1 用户访问控制

“全国煤炭资源网络信息系统”是基于 B/S 的大型网络数据发布平台, 是全国范围的、面向大众的复杂网络应用系统。不同的用户必须具有不同的工作角色和权限级别, 不同的权限所对应的获取、处理信息的需求和权限也不相同。为了区分不同用

收稿日期: 2008-04-11

作者简介: 杨圣伟(1983—), 男, 湖北丹江口人, 硕士研究生, 主要研究方向为国土资源信息化。

户和不同权限，必须对用户的访问权限进行控制。

“全国煤炭资源网络信息系统”所承载的数据属于全国性矿产数据，具有一定的涉密性，而且数据量巨大、用户人数众多、处理事务繁多，这就要求它既能够满足用户合理地访问系统，又要系统管理人员能够统一、高效、灵活地进行访问控制。传统的访问控制机制，如自主访问控制(DAC，Discretionary Access Control)、强制访问控制(MAC，Mandatory Access Control)等已远不能满足访问控制的要求，而 20 世纪 90 年代以来发展起来的基于角色的访问控制(RBAC，Role Based Access Control)技术，则可以减少授权管理的复杂度，降低管理开销，提高访问控制的安全性^[2]。

1.2 用户访问控制的实现

“全国煤炭资源网络信息系统”中，用户访问权限控制是采用基于角色的访问控制(RBAC，Role Based Access Control)技术，它使用“用户 - 角色 - 权限”3 层模式对用户进行访问权限控制。

用户隶属于角色，用户和角色之间是多对一的关系，一个用户只能隶属于一个角色，一个角色可以有多个用户。角色对应权限，角色和权限采用多对多的关系，一个角色可能有多个权限，一个权限也可能赋予多个角色。

考虑到系统的开放性，系统对用户角色进行了多层分级，使得用户权限相应细分。本系统使用的角色有：

超级管理员(仅 1 人) 开发用户；

系统管理员(1 人或数人) 拥有所有权限，包括可以进行用户管理，但不能对超级管理员的账号进行管理，其具有负责管理系统，更新数据等职责；

高级用户 内部注册用户，其没有用户管理权限，可以操作数据，具有对数据进行查询、下载等权限；

一般用户 外部注册用户，其可以浏览、查询数据；

游客 未注册用户，其可以浏览系统结构和部分元数据，但不能浏览数据。

表 1 为系统用户控制采用的数据字典。

系统中，使用 role 和 User_limits 字段规定用户角色和权限，引入 Login_time 和 End_time 两个字段记录用户登录时间和离开时间，使用 Total_time 字段规定用户在多长时间无操作的情况下断开数据库连接，减少数据库联接数量，控制数据安全。使用 province 字段记录用户所属省份(部门)，限定一般注册用户只能查看本省(本部门)数据，而没有权限查

看全部数据。

表 1 用户控制使用的数据字典
Table 1 Data dictionary on user control

列名	数据类型	数据长度	允许为空
username	char	8	否
User_limits	varchar	50	是
password	char	10	否
Login_time	datetime	8	否
End_time	datetime	8	否
Total_time	float	8	是
province	char	50	否
User_photo	binary	256	是
Role	varchar	50	否

2 基于数据字典的语义翻译实现

在数据库建立时，大多对数据模型中的数据项进行了编码工作，使得数据入库后方便操作，但是在数据发布系统中，如果用户看到的都是编码，那么该数据项的意义就无法明确，为了准确地表达数据项的意义，必须对数据编码进行规格化翻译。

在 Microsoft. NET 平台中，大都采用 Grid View 模板的 Header Text 属性逐个修改。但是因为地质数据与其他数据相比，具有多源、多量、多类、多元、多维的特点^[3]，逐个修改整个系统的数据项编码是不可行的。系统采用数据字典技术对数据进行翻译，使得数据所代表的语义清晰明确。

表 2 为系统对数据进行语义翻译所采用的数据字典。其中，table_name 字段为使用该数据项的表名称，field_name 为数据项编码，field_capt 为数据项的中文语义，field_type、field_len 填写数据项的类型和长度，field_defa 和 field_rule 记录数据项的默认值和填写规则。

表 2 语义翻译和界面定制使用的数据字典
Table 2 Data dictionary on data meaning translating and dynamic customization

列名	数据类型	长度	允许为空
table_name	char	16	否
field_name	char	24	否
field_type	char	1	是
field_len	float	8	是
field_defa	char	16	是
field_rule	char	10	是
field_capt	char	100	否
field_from	char	6	是
field_input	char	24	是
twice_mean	char	10	是

3 利用数据字典实现用户定制界面显示

“全国煤炭资源网络信息系统”涉及的数据量大，数据结构复杂，为了完整表述某一实体的全部属性，一张表往往具有十几甚至几十个数据项。一方面，用户所关心的可能仅仅是这张表的若干数据项，而不是整张表，故不需要显示整张表，这样既可以突出用户所关心的内容，也可以减少网络数据流量；另一方面，因为数据量巨大，如果在客户端全部显示，则界面宽度将很大，这在目前主流的浏览器配置下，非常不利于用户使用。鉴于以上两个原因，在浏览器端，系统便采用用户定制显示的方式来显示数据。

一般数据库的界面和底层数据库之间是一一对应关系，即一个界面对应一个数据表，如果要实现定制，必须打破数据表和界面的一一对应关系。在数据表和界面之间加入数据字典，使得界面和数据字典相对应。利用数据字典记录数据项的来源和数据项的其他基本信息，使得数据表和显示界面分离。在页面显示端，用户选择需要显示的数据项(图 1)，将选择结果反馈给服务器端。根据用户选择显示的

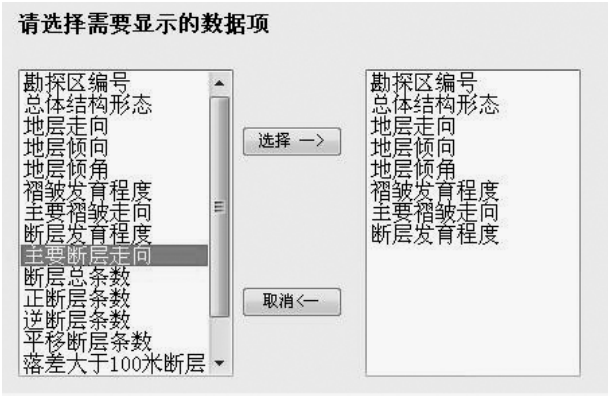


图 1 选择显示的数据项

Fig.1 Choose the data to display

数据项，服务器从对应数据表中提取数据项，结果显示在浏览器端(图 2)。

完成动态显示的数据字典参见表 2。其中，field_from 记录了字段的来源，即用来记录用户需要查看该数据时，系统从哪个表提取该数据项；field_input 记录了当用户在浏览器端写入数据时，该数据项要写入哪些表。field_from 和 field_input 还可以用来完成一致性控制；twice_mean 记录该字段的其他意义，还可以完成多意数据项的翻译。

勘探区编号	总体构造形态	地层走向	地层倾向	地层倾角	褶皱发育程度	主要褶皱走向	断层发育程度
2303000404	8	北北东	北西	15-25	3	北北东	3
2303000303	7,8			15-25	3	SN	2
1550000200	6			7-12	3	北北东	2
2309000200	3	N20-75	N10-25	20			2
2307000200	1						2
2310000100	1				1		1
2307000100	5,6			3-5	2	北北东	1
2308000400	1	水平			1		1
2308000300	1			0-8	1		1
2308000200	1			0-8	1		1

图 2 浏览器端的结果显示

Fig.2 Result of dynamic customization

4 结 语

以数据字典在“用户访问控制”、“数据语义规格化”、“显示页面动态定制”等方面的应用为例，探讨数据字典技术在煤炭数据发布系统中的具体应用。通过使用数据字典技术，完成了多用户、多角色和多权限的用户访问控制功能，实现了多元、多量、多类、多元和多维数据的语义翻译和基于用户的页面动态显示定制功能，保证了系统的安全性和功能的完善性。

参考文献

[1] 汪新庆, 刘刚, 袁艳斌, 等. 地质矿产术语分类代码在地矿点源信息系统中的应用[J]. 地球科学-中国地质大学学报, 1999 24(5): 529-532.

[2] 周小为, 何斌. 基于角色的授权与访问控制的研究[J]. 计算机与信息技术, 2006(10): 86-88.

[3] 吴冲龙, 汪新庆. 地质矿产点源信息系统设计原理及应用[M]. 武汉: 中国地质大学出版社, 1996.