MULTI LEVEL AND MULTI LANGUAGE DATABASE CODING SYSTEM FOR DATA NORMALIZATION OF VETERINARY DATABASES

Mohd Termizi Ghazali  
termizighazali@toki.waseda.jp

Achmad Rully  
arullly@ruri.waseda.jp

Yoshiyori Urano  
yurano@waseda.jp

Global Information & Telecommunication Studies  
29-7 Waseda University Bldg., 1-3-10 Nishi-Waseda, Shinjuku-ku,  
Tokyo 169-0051, Japan  
Tel: +81-3-5286-9847  
Fax: +81-3-5286-3832

ABSTRACT

The usage of code reference database table for data normalization has been extended as an interface for storing and displaying of coded data in a database application system. This concept has been applied in the development of alternative veterinary database application for managing event-based data growth. Relational Database Management System (RDBMS) data structure has been deployed for the creation of active and code reference database tables.

The database designing strategy is focusing on the data uniformity, normalization, and reduction of record size for every active database table in the system. Data coding is seen as an alternative to achieve the objective. Integration of second level data normalization (2NF) and data coding showed positive result to improve database table handling in the study. The concept of data coding is deploying 4 fields or columns in the development of code reference database table's schema for maintaining multi level associated data of active database tables in the application system.

In this study the concept is applied into veterinary databases by using 2 active database tables for maintaining event-based data collection and 1 code reference database table for coded data translation. The database tables are event-based data collection, animal background information and coded field data information, respectively. It shows positive result in reducing record size of active database table and number of table required by the system.

Experiment of the study showed that maximum record size reduction of active database tables is 45%. However the reduction percentage is vary according to numbers of fields that can be coded and also original field length of the database tables. Further more, number of database tables required in maintaining the system also reduced from 14 tables into 3 tables. It is achieved by deploying 2NF instead of 3NF in database designing phase. Single and multi level coding concept that are maintained by code reference database table becomes main factor to the possibility of adoption 2NF in the database design. Multi language translation could be integrated into the coding concept based on the code translation algorithm.

There is a plan in the near future to study possibility of this concept to be expanded in distributed system environment for integrated database development. The concept would be applied to the Department of Veterinary Services of Malaysia's (DVS) for data warehouse development. The basic principle for the expansion is storing code reference database table in the workstation whereas the active database will be in the database server. The main focus is measuring data transmission rate in the application system.

Keywords
Data normalization, database coding concept, RDBMS

1. INTRODUCTION

In this paper the intention is to highlight an alternative database designing model and application system that could be applied for the Department of Veterinary Services of Malaysia (DVS) [1]. In addition, data normalization concept is being highlighted in the designing and developing of the small components that could be expendable to the whole system in the near future.

Designing and developing the alternative model should suite the DVS database management system in resolving some of the constraints due to the older version of the database application system that could not operate well in the multi-platform environment. Further more the older version of database structure could not optimize the full capability of advanced information technology hardware and concurrent requirement of data capture.

In addition, tremendous incremental of data growth in the system over 200 operational centers operated nation wide [1] could be regarded as another factor why the alternative model should be proposed. The daily event-based data capture is focusing on the animal treatment, production and reproduction system. It is involving species of animal, type of treatment given, number of animal treated, disease occurrence, location, farmer or client information, and etcetera. Besides, image data could be necessary to be stored in the system for the case reference in the future. The usage of accumulated data in the department could be varies such as planning, disease surveillance and monitoring, reporting, and etcetera [2].

These mean that the new database platform should support and yet be integrated with various database formats for future expansion in developing of data warehouse. Thus RDBMS that could support metadata entry and distributed operability would be favorable for the database platform. However in this study database designing is given priority and will be discuss in detail since it is crucial part in developing of any database oriented application system [3, 4].

Therefore the main activity of study is focusing on possibility of reducing record size of database table during normalization phase. One of the approaches is introducing coded data incorporating with normalization to enhance database integration of the application system.

The higher data normalization are, the better structured it would have. How ever the system could becomes more complex and would effecting processing performance [3, 4]. Besides,