Agricultural Field Trials – Today and Tomorrow, 166-169 (2007)

A systematic database solution as an efficient way of modern management of grassland and forage crop experiments

Reinhard Resch[1] and Wilhelm Graiss[1]

Introduction

Research on grassland and forage crops has today (2007) a tradition of 60 years at the HBLFA Raumberg-Gumpenstein (Agricultural Research & Education Centre). Beginning with the late 1980ies, data management has been changed step by step from hard copy documentation to electronic database systems. Grassland and forage crop experiments mostly deal with different objectives regarding utilisation, fertilisation, botanical composition, plant breeding, testing of seed mixtures and cultivars or plant protection and therefore complex schemes and database structures are required for a modern trial management and multivariate data analysis.

Each department and each experiment has individual database requirements, but there is no standard software for management of complex dataset in grassland experiments so far. To solve fundamental problems occurring in the management of data from grassland experiments, it is necessary to design individual database models, schemes and table structures. We therefore devised a strategy, solving individual problems within a general useable database.

Materials and methods

A SQL (Structured Query Language) database, which has been a well established and sustainable standard for more than 30 years, was chosen at the HBLFA Raumberg-Gumpenstein. Relational databases are a persistent system which represents a lot of different functional tools. SQL data tables can contain an enormous capacity of datasets and there are no table space problems in data saving if you want to save a million of datasets or more. The input of different variables in one table is not limited with 256 fields (columns), and a SQL data table can include approximately 2000 variable names. The most common SQL software package ORACLE (Heitsiek, 2006) was too expensive. Therefore, we chose the open source software PostgreSQL (PostgreSQL Global development group, 2003; Hartwig, 2001).

Before we started to use PostgreSQL we have collected a lot of experience with dbase and later on with Microsoft Access. The advantage of Microsoft Access is the simple control of program routines with Visual Basic and the well-functioning connection with different types of data tables like dbase, Excel, etc. Now, with the implementation of PostgreSQL, we use Microsoft Access as front end, mainly the features of forms and reports. The connection of PostgreSQL data tables with Microsoft Access is easily possible, as there is an ODBC (Open DataBase Connectivity) driver freely available (http://pgfoundry.org/projects/psqlodbc/, last visit May, 21st, 2007).

[1]Federal Research & Education Centre Raumberg-Gumpenstein
Institute of Plant Production and Cultural Landscape
A-8952 Irdning, Altirdning 11
Tel.: +43 (0)3682 22451, Fax: +43 (0)3682 22451-210
E-Mail: reinhard.resch@raumberg-gumpenstein.at, wilhelm.graiss@raumberg-gumpenstein.at

Another aim of our database system is the connection to the intranet and the World Wide Web (www). The importance of interdisciplinary projects with partners of different countries is developing and we need common data platforms for an efficient data management and exchange. At the HBLFA Raumberg-Gumpenstein we use the language PHP (http://www.php.net, last visit May, 21st, 2007) to build the optimal communication between client and server.

Our trial management database is administrated by three persons, who are well established in agriculture and have a high expertise in plant production and in scientific data management.

Results and discussion

At the beginning of the design of a good database system, detailed knowledge of the technical domains and its structures is required. The aim for us was a well adapted system for the Institute of Plant Production and Culture Landscape at the HBLFA Raumberg-Gumpenstein. This institute includes three divisions and four departments with different responsibilities. The highest priority was a database system for field experiments, especially forage crops.

A database system has to maintain tables with data from different sources and must connect them with one or more mother tables, which include basic information of a functional domain like experiment number, experiment name, date of beginning, plot size, account number, etc. Each data table needs firstly a primary key to identify each dataset and secondly also one or more defined foreign key variables. The foreign key is essential to combine two or more tables together. The secret of a successful database system is the combination of small structured data tables by those defined keys.

There are a lot of advantages if more than one powerful software package is used and some problems can successfully be solved by this strategy. It is often necessary to program calculation routines or difficult procedures for data analysis which can efficiently be done by the use of SQL, VISUAL BASIC or PHP, depending on the specific purpose. It is very important that the used language is simple and can easily be learned by own technical experts at the different institutes.

The database system for grassland and forage crop experiments at the HBLFA Raumberg-Gumpenstein is based on a model that looks like a Greek temple (Fig. 1). The foundation is a data table that includes important keys and the main information about experiments of different departments. There are six supporting pillars with grouped functional data contents in the “database building”. The roof of the building stands for the global experiment documentation and simultaneously for the quality management function of the database. Forage crop experiments often analyse different species and/or cultivars of plants or different types of seed mixtures. Every variant of a field trial has to be defined and saved in an adequate data table. All of the experimental definition details of variants are part of the first database column. The sources of these tables serve a documentary function.

The second data domain includes the plot archive of all experimental stations also weather and geographic data tables (coordinates, parcel codes, commune numbers, production regions, etc.). Spatial data is necessary to combine experimental information with a GIS (Geographic Information System).

In the third column of the system there are all of the tables with measured experimental data like yields, analysis data of soil, forage and fertilizers, botanical composition, etc. The connection of electronic equipment like scales, handheld PC´s, combined harvesters or analytical instruments via an interface to the database system is very effective to exchange and combine measured data. Peripheral data streams can be collected in different ways for example by direct transfer via wired or wireless LAN (intranet), data logging or data saving to external memory. It is essential to have well defined, open and documented interfaces in combination with easy-to-use programmed tools for a speedy import of data into the data base system without any mistakes and losses.

Validation of experimental data is very important to get best quality values for the statistical data analysis. It makes sense to combine a quality management with the database system to reduce sources of error. The more functional tables are included in the database the more combined and complex questions can be asked and answered. The output is always a matrix of data that can be exported into a different data format. The import of those data files in a statistical software package (SAS, SPSS, Statgraphics, R, etc.) is mostly well supported.

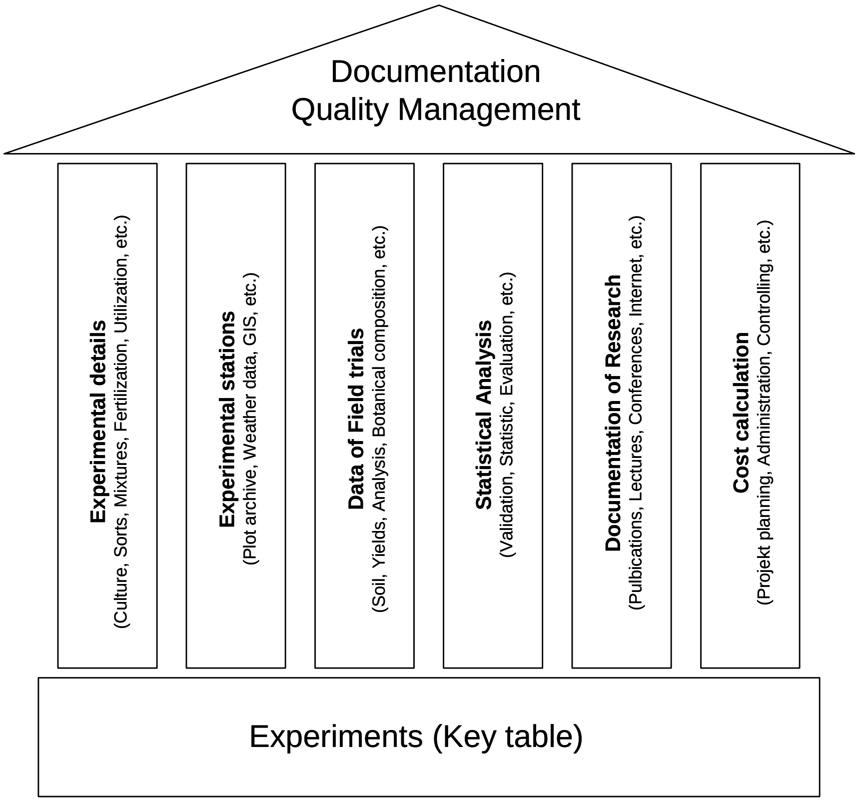

Fig. 1: Database model for grassland and forage crop experiments implemented at the HBLFA Raumberg-Gumpenstein

At the HBLFA Raumberg-Gumpenstein a information system for the research work (publications) was developed in-house on the basis of PostgreSQL in combination with an intranet application written in PHP. This “Research Documentation Base” (Forschungsdokumentation - FoDok) includes citations of publications, lectures, conferences, etc. with corresponding PDF files. It is possible to create required up to date reports for the ministry of agriculture and forestry water management and environment in a few minutes. Another use case of this information system is the web site of our research and education centre. Any visitor of http://www.raumberg-gumpenstein.at can retrieve results and download files from experiments of each department in an efficient and quick way. The synchronisation works via automatic tasks during the night, the data is fed into the system once and is used in several ways (reporting to the superior authority, content for the web site...) without any further manual processing.

Cost calculation is also a task for a modern database system. Every defined experimental project has an account number. The working time from every person for each experiment has to be recorded in a tabular form. It is quite simple to calculate the time with actual personal costs and you get the labour costs of every defined experiment. This database tool is very important for project management and administration. It is a controlling instrument to analyse the real costs of experiments, departments and institutions and to be able to compare them with the planned costs.

If more than one person is working with a database system it is necessary to distribute individual authorizations for reading, writing and deleting of data. The maintenance of data authorization has to be in the hands of the system administrator in cooperation with the owner of the data. Good SQL databases like PostgreSQL (open source) or ORACLE provide a well-working authorization system for granting individual rights for each user.

Summary

The Institute of Plant Production and Culture Landscape of the HBLFA Raumberg-Gumpenstein is aiming at a modern trial management for grassland and forage crops on the basis of a full electronic documentation combined with quality management and cost calculation. All recorded data of three divisions and four departments should be transferred into a sustainable database system to get best information as quickly as possible. The database system is grouped in six functional data contents (experimental details, experimental stations, data of field trials, statistical analysis, documentation of research and cost calculation) based on a key table with essential experiment data. If all validated experiment data are collected and saved in the efficiently administrated database system, it should be possible to get the requested results in optimal quality by “just pressing a button”.

The development of an individual, well-adapted database system for your own needs is a hard and stony way and requires a lot of patience, but is a basic and essential investment for a successful scientific work.

References

Hartwig, J., 2001: PostgreSQL Professionell und praxisnah. Verlag Addison-Wesley, ISBN 3-8273-1860-2

Heitsiek, S., 2006: Oracle Database 10g Express Edition. Der schnelle Einstieg. Verlag mitp, ISBN 3-8266-1670-7

PostgreSQL Global Development Group, 2003: PostgreSQL Das offizielle Handbuch. Verlag mitp, ISBN 3-8266-1337-6